

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 22:46:32 ; Search time 504 Seconds
(without alignments)

6024.609 Million cell updates/sec

Title: US-10-761-033-1_COPY_1_1316

Perfect score: 1316

Sequence: 1 atgacaggagcaagaggaa.....gtgataaaatcgattgggg 1316

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7673375 seqs, 115364844 residues

Total number of hits satisfying chosen parameters: 15346750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA New:*

- 1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB_seq.*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB_seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB_seq.*
- 4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB_seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB_seq.*
- 6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB_seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB_seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB_seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB_seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------------------|--------------------|
| 1 | 214.2 | 16.3 | 2829 | 7 US-10-505-263-88 | Sequence 88, Appl |
| 2 | 203.2 | 15.4 | 3635 | 7 US-10-505-263-80 | Sequence 80, Appl |
| 3 | 184.6 | 14.0 | 3200 | 7 US-10-505-263-86 | Sequence 86, Appl |
| 4 | 178.2 | 13.5 | 197096 | 12 US-11-121-086-107 | Sequence 107, Appl |
| 5 | 177. | 13.4 | 2640 | 7 US-10-505-263-82 | Sequence 82, Appl |
| 6 | 176.6 | 13.4 | 3367 | 7 US-10-505-263-11 | Sequence 11, Appl |
| 7 | 172 | 13.1 | 4815 | 7 US-10-537-002-56 | Sequence 56, Appl |
| 8 | 168.8 | 12.8 | 2654 | 7 US-10-505-263-1 | Sequence 1, Appl |
| 9 | 168.8 | 12.8 | 2748 | 7 US-10-505-263-3 | Sequence 3, Appl |
| 10 | 166.4 | 12.6 | 2279 | 7 US-10-505-263-90 | Sequence 90, Appl |
| 11 | 164.2 | 12.5 | 2612 | 7 US-10-505-263-5 | Sequence 5, Appl |
| 12 | 164.2 | 12.5 | 2750 | 7 US-10-505-263-7 | Sequence 7, Appl |
| 13 | 157.8 | 12.0 | 2457 | 7 US-10-505-263-84 | Sequence 84, Appl |
| 14 | 109 | 8.3 | 2871 | 7 US-10-505-263-9 | Sequence 9, Appl |
| 15 | 48.6 | 3.7 | 2290 | 9 US-11-096-568A-33954 | Sequence 33954, A |
| 16 | 44.8 | 3.4 | 2298 | 9 US-11-096-568A-31695 | Sequence 31695, A |
| 17 | 44 | 3.3 | 10490 | 12 US-11-011-332A-87 | Sequence 87, Appl |
| 18 | 43.4 | 3.3 | 10490 | 12 US-11-011-332A-101 | Sequence 101, Appl |
| 19 | 43.2 | 3.3 | 610 | 6 US-09-925-065A-422255 | Sequence 422255, |
| 20 | 43.2 | 3.3 | 610 | 6 US-09-925-065A-422256 | Sequence 422256, |

```

21 43 3.3 1452 8 US-10-793-626-3131 Sequence 3131, Ap
22 43 3.3 3926 8 US-10-793-626-4300 Sequence 4300, Ap
23 43 3.3 4189 8 US-10-793-626-3898 Sequence 3898, Ap
24 42.8 3.3 610 6 US-09-925-065A-422254 Sequence 422254,
25 41.6 3.2 568 6 US-09-925-065A-264855 Sequence 264855,
26 40.4 3.1 2417 12 US-11-010-239-47 Sequence 47, Appl
27 39.8 3.0 1004 8 US-10-793-626-4406 Sequence 4406, Ap
28 39.6 3.0 503 6 US-09-925-065A-558945 Sequence 558945,
29 39.6 3.0 548 6 US-09-925-065A-558944 Sequence 558944,
30 39.6 3.0 888 7 US-10-932-182A-3583 Sequence 3583, Ap
31 39.6 3.0 888 7 US-10-932-182A-3583 Sequence 3583, Ap
32 39.6 3.0 1695 12 US-11-098-686-9652 Sequence 9652, Ap
33 39.6 3.0 1457619 12 US-11-098-686-8739 Sequence 8739, Ap
34 39 3.0 643 6 US-09-925-065A-769517 Sequence 769517,
35 38.8 2.9 2685 7 US-10-932-182A-1726 Sequence 1726, Ap
36 38.8 2.9 2685 7 US-10-932-182A-1726 Sequence 1726, Ap
37 38.2 2.9 640 6 US-09-925-065A-623209 Sequence 623209,
38 38 2.9 770 7 US-10-349-331-225 Sequence 225, App
39 37.8 2.9 890 12 US-11-136-527-1445 Sequence 1445, Ap
40 37.8 2.9 890 12 US-11-136-527-5541 Sequence 5541, Ap
41 37.6 2.9 148220 12 US-11-121-086-90 Sequence 90, Appl
42 37.6 2.9 179597 12 US-11-121-086-91 Sequence 91, Appl
43 37.2 2.8 524 6 US-09-925-065A-405101 Sequence 405101,
44 37 2.8 148935 8 US-10-995-561-13308 Sequence 13308, A
45 37 2.8 1082144 12 US-11-117-187-211 Sequence 211, App

```

ALIGNMENTS

RESULT 1

```

US-10-505-263-88
; Sequence 88, Application US/10505263
; Publication No. US20060014940A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Case Western Reserve University
; APPLICANT: The Brigham and Women's Hospital, Inc.
; APPLICANT: Mount, David B
; APPLICANT: Romero, Michael
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF SLC26A6, and SLC26A2
; TITLE OF INVENTION: ANION EXCHANGERS
; FILE REFERENCE: 1242/50/2 PCT/US
; CURRENT APPLICATION NUMBER: US/10/505,263
; PRIOR FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: US 60/360,275
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US03/06469
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 88
; LENGTH: 2829
; TYPE: DNA
; ORGANISM: Xenopus laevis
US-10-505-263-88

```

Query Match 16.3%; Score 214.2; DB 7; Length 2829;

Best Local Similarity 49.1%; Pred. No. 2.8e-46;
Matches 598; Conservative 0; Mismatches 613; Indels 6; Gaps 1;

```

QY 95 TGCCCATTTTGGATTTGGGCACCACTTACATCTGAAAGAAAACCTTGTTCAGACACTG 154
DB 521 TTCCAATTTCTGGTTGGTTTACCGGTATCCAGTAAGGAATGGCTGCTCGAGACATAG 580
QY 155 TGCTGGGATTAATGTGGCAGTTCACAGGTGACCAAGGATGGCCTTTGCTTCTCT 214
DB 581 TATCTGGTCTTAGTGTGGCATCATTCAGCTCCCAAGGTTTACTGTTATGCTTTACTGG 640
QY 215 CATCTGTCACCCAGTGTTCGTTTATATGGTCTCTCTTCTCCGCCCAATATTTATGCCA 274
DB 641 CTGAGATTCCACCCGGTGTGGATTGTTATTCCTCATTTTCCCTGTACTTCTCTATGCCA 700

```

275 TATTTGGAAATGGGACATCATGTTGGCCACAGGACACTTTGGCTTGACATCTCTTAATATACAG 334
701 TATTTGGGACTTCAAGACATATTTCCCTGGAACTTTGCTGTGATTTCTGTGATGTTG 760
335 CCAAGCCGTTGGAACGGATTTCCCTCAGACATGCAGAACTCCACACAGAGTAACA 394
761 GTAGTGTGACTGAATCACTGGTGCC-----GTCAAGAACTTACAGGTGTACTCTGGAATG 814
395 CAAGCGTCTGGCTTATCCGACTTTGAAATGCAAAAGATCCACGTTGCTGCGACAGTTT 454
815 AATCAGTGAATGATATAGCAGCCAGAGCAATGACAGAGTTGAGTGGCATCGGCCCTGA 874
455 CTTCTTGGAGGTGATTCAGTGGGCAATGTTTGTGCTGCAACTGGGCGAGTGCCACAT 514
875 CATTCCTAGTTGGACTTTTTCAGATTATGTTGGGTCTGTCGCAAGTTGGCTTGTGTA 934
515 TTGTGTCACAGAGCTGTGATCAGCGCAATGACAACTGGGGCTGCCACCCATCTGGTGA 574
935 CTTACCTCTCTGAACCCCTTGATCCGAGGCTACACAGTGTGCTGCAATCCATGTCACTG 994
575 CTTCAAGTCAAAATATCTTTGGGAATGAAATGCCATATATATCCGGACCACTTGGAT 634
995 TGTCTCAGATGAAGTGTCTTAGGCGTGAGATCAGCAAGAGATCATCATCTGCTC 1054
635 TCTTTTATATTTATGATATGTTTGTGAAATCATCAAGTCTGTGCGACTGGAAGCATTGC 694
1055 TAAATTTATGCAATCTGTGCAAGTCTCTGCGCAAGTCTCTGAGACAAATATTTGCTTCCCTTC 1114
695 TTTTATCCTCTGAGCAATGTGGTCTGTTCTGTTTAAAGAGCTGAATGAACAGTTTA 754
1115 TAAATGGGTGCACTCCCATCACTGTCTGTTTCTGTTTAAAGTTTCTTAAACGACAAGTACA 1174
755 AAAGGAAATTAAGTGTCTCTCTGTGATGTTAGTTTTCATATTTGCTGCATCATTTG 814
1175 GCTCAAGATCGTATGCCCATACCAATCGAATCATCACTGATGTTAGTCCACTGGGA 1234
815 CTTGTTATTTGCAACATATGAAACACATATGGAATAGAAATGTTGTTGTTGTTTCCAC 874
1235 TATCATATGTTGCCAGTCTCCACCAAGTCTATGAGTGGACATTTGTAGGGAGATCCAA 1294
875 AAGNAATCCCTCACTAGAGCTCCCGATGAACATCTCTCTGCGGTGATCACTGAG 934
1295 CTGGATGAAGCTCCAATGTTACCAACACAAATATTTTTCGAAGAGTGGTGGCAATG 1354
935 CTTTCGGAGTGGCACTTGTAGGCTATGTTGGCTCACTGGCTCTTGTCTCAAGGATCTGCCA 994
1355 CTTTGGCATCGTGTGTTGTGCTATGCTTTACATCTCCCTGCTAAGATGTTGGTG 1414
995 AAAAATTCAAATATTCATATGATGACAAACAGGAATTTTGGCCCATGGCTCAGCAATA 1054
1415 TTAACATGGATACAAATATGACAGCAATCAGGAATCAGTCCCTTGGGGCTGAGTAAT 1474
1055 TAGTTTCTTCAATTTTCTTCTGATACCAAGTGTCTGCTCCATCGGAGGAGCGCTGCC 1114
1475 CAATAGGAAGCTTCTTTCAGTGTCTTCAATTTGGGACAGCCATGTCAAGAGTCTGGTGC 1534
1115 TGTACAGCACAGGAGGAGACACAGGTGGCTGTCTTAATATCTTGTGCAATTTTCTCTCTTA 1174
1535 AAGAAAGCACAGGAGGAGACAGTCAAGTACGAGTGTGCTCTCTCTTGTGCTATCTT 1594
1175 TAGTCATCTATGCAATAGGACCTTTGCTTTTACTGGCTGCCCATGTGTCTCTTCCAGACA 1234
1595 TTATCATTTAAAGCTGGGAGCTGTTTGAACCCCTTACCAAAAGCCATCTATGCTCGAG 1654
1235 TTATGTTGGGAGTGAAGGATGCTAATACAGTTCGGAGATTTTAAAGAAATTTTGA 1294
1655 TTGTTGTAGTCAATCTCAAGAGTATTTTACAGAGTGTACTGATGTGCCAATGTTATGGA 1714
1295 ATGTGGATAAATCGAT 1311
1715 GATCAACAAATTTGAT 1731

RESULT 2
US-10-505-263-80
; Sequence 80, Application US/10505263
; Publication No. US20060014940A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Case Western Reserve University
; APPLICANT: The Brigham and Women's Hospital, Inc.
; APPLICANT: Mount, David B
; APPLICANT: Romero, Michael
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF SLC26A6, SLC26A1, and SLC26A2
; TITLE OF INVENTION: ANION EXCHANGERS
; FILE REFERENCE: 1242/50/2 PCT/US
; CURRENT APPLICATION NUMBER: US/10/505,263
; CURRENT FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: US 60/360,275
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US03/06469
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 80
; LENGTH: 3635
; TYPE: DNA
; ORGANISM: Xenopus laevis
US-10-505-263-80

Query Match 15.4%; Score 203.2; DB 7; Length 3635;
Best Local Similarity 48.5%; Pred. No. 2.6e-43;
Matches 596; Conservative 0; Mismatches 623; Indels 9; Gaps 1;

Qy 94 CTGCCCAATTTGGATTTGGGACCAACACATTAACAATCTGAAAGAAACTTGTCTTCAGACACT 153
Db 241 CTCCTATTCTGAATTTGTTACCAATAACAGATGGAAGATGTTTTTAAGTGACCTG 300
Qy 154 GTGCTGGGATAATGTTGGCAGTTTCAACAGGTGACCAAGGATTTGGCTTTGCTGTTCTC 213
Db 301 ATCTCTGGAGTCAAGCAGGCTTGTGGTACCTTTGCAAGGCTTGGCAATTTGCAATGCTT 360
Qy 214 TCATCTGTGACCCAGGTTTGGTTTATATAGGCTCTCTGTTTCTGCTGCATATTTATGCC 273
Db 361 GCTGAGTCCCTGTTGGATATGGCTCTACTCTTCTTTTCTCTATCTTGTGACTTATTTT 420
Qy 274 ATATTTGGAATGGACATCATGTTGGCCAGGCACTTTGCTTTGACATCTCTTAATATCA 333
Db 421 TTCTGGGACATCAAAACACATTTCTGTGGACCAATTCCTCAGTGGTCAAGTTGATGTTG 480
Qy 334 GCCAACCCCTGGAAACGATTTGCTTCCCTCAGAACATGACAGATCT-----CACACA 384
Db 481 GGATCAGTTTCTGTCTATTTGCTCCAGATGAAAAATTTGCAATATATTAGGCAACTTACA 540
Qy 385 CAGAGTAACAACAGCGTCTGGGCTTATCCGACTTTTGAATGCAAGATCCAGTTGCT 444
Db 541 GGACTGAACAAAACAATAATAGACACAGTAGAAGATGAGCCAGAGTTGCCGTTTCA 600
Qy 445 GCAGCAGTTTCTTCTTGGGAGGTGATTTCAAGTGGCCATGTTTGTGCTGCAACTGGGC 504
Db 601 GGAACACTCAGCTTCTTAATTGGNAATATACAGCTTGCACCTTGGCGTATTTCAAAATGGC 660
Qy 505 AGTGCCACATTTGTGGTCAACAGGCTGTGATCAGCGCAATGACAACTGGGGCTGGCAACC 564
Db 661 TTTATTATCAGATATTTGGCAGATCCATTAGTTGGTGTGTTTACTACAGCTGCAAGCTTC 720
Qy 565 CATGTGTGACTTTCACAAGTCAAAATATCTCTTGGGAATGAAATGCCATATATATCCGA 624
Db 721 CAAGTGTGTTTTCACAAATCAAACTAAGTGTCTCAATGTCCCAACAAAGATCTACACGA 780
Qy 625 CCACTTGGATCTTTTATATTTATGCTATGTTTGTGAAACATCAAGTCTGTGCGACTG 684
Db 781 GTACTCTCTATCATATATACAAATAATAGATATCTTTACCAACATTTGCAAAACCAACAT 840
Qy 685 GAAGCATTTGCTTTTATCTCTTGTGAGCATTTGGTCTCTGTTTCTTGTGTTAAAGAGCTGAAT 744

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 22:44:50 ; Search time 1196 Seconds
(without alignments)
9099.081 Million cell updates/sec

Title: US-10-761-033-1_COPY_1_1316

Perfect score: 1316
Sequence: 1 atgacaggagcaagaggaa.....gtgataaaatcgattgggg 1316

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA_Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-----------------------|-------------------|
| 1 | 1316 | 100.0 | 1797 | 3 US-09-875-811-11 | Sequence 11, Appl |
| 2 | 1316 | 100.0 | 1797 | 7 US-10-761-033-11 | Sequence 11, Appl |
| 3 | 1316 | 100.0 | 1818 | 3 US-09-875-811-7 | Sequence 7, Appl |
| 4 | 1316 | 100.0 | 1818 | 7 US-10-761-033-7 | Sequence 3, Appl |
| 5 | 1316 | 100.0 | 1866 | 3 US-09-875-811-3 | Sequence 3, Appl |
| 6 | 1316 | 100.0 | 1866 | 7 US-10-761-033-3 | Sequence 9, Appl |
| 7 | 1316 | 100.0 | 1971 | 3 US-09-875-811-9 | Sequence 9, Appl |
| 8 | 1316 | 100.0 | 1971 | 7 US-10-761-033-9 | Sequence 1, Appl |
| 9 | 1316 | 100.0 | 1971 | 9 US-10-505-316-1 | Sequence 1, Appl |
| 10 | 1316 | 100.0 | 1992 | 3 US-09-875-811-5 | Sequence 5, Appl |
| 11 | 1316 | 100.0 | 1992 | 7 US-10-761-033-5 | Sequence 5, Appl |
| 12 | 1316 | 100.0 | 2040 | 3 US-09-875-811-1 | Sequence 1, Appl |
| 13 | 1316 | 100.0 | 2040 | 7 US-10-761-033-1 | Sequence 1, Appl |
| 14 | 1316 | 100.0 | 2345 | 5 US-10-222-009-12 | Sequence 12, Appl |
| 15 | 1316 | 100.0 | 2420 | 3 US-09-875-811-13 | Sequence 13, Appl |
| 16 | 1316 | 100.0 | 2420 | 7 US-10-761-033-13 | Sequence 13, Appl |
| 17 | 1314.4 | 99.9 | 1971 | 3 US-09-942-446-3 | Sequence 3, Appl |
| 18 | 1314.4 | 99.9 | 2631 | 3 US-09-814-353-20997 | Sequence 20997, A |
| 19 | 1314.4 | 99.9 | 2886 | 5 US-10-222-009-11 | Sequence 11, Appl |
| 20 | 1314.4 | 99.9 | 5267 | 3 US-09-942-446-1 | Sequence 1, Appl |
| 21 | 1166.2 | 88.6 | 2848 | 7 US-10-415-378-24 | Sequence 24, Appl |
| 22 | 1068 | 81.2 | 2289 | 9 US-10-505-316-3 | Sequence 3, Appl |
| 23 | 514.4 | 39.1 | 1296 | 9 US-10-450-763-6290 | Sequence 6290, Ap |

ALIGNMENTS

RESULT 1

; Sequence 11, Application US/09875811
; Patent No. US20020032321A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; TITLE OF INVENTION: No. US20020032321A1 Human Transporter Proteins and Polynucleotic
; FILE REFERENCE: LEX-0186-USA
; CURRENT APPLICATION NUMBER: US/09/875,811
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/210,045
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-875-811-11

| | | | | |
|-----------------------|--------------|---|------------|--------------|
| Query Match | 100.0% | Score 1316; | DB 3; | Length 1797; |
| Best Local Similarity | 100.0% | Pred. No. 0; | | |
| Matches 1316; | Conservative | 0; | Mismatches | 0; |
| | | | Indels | 0; |
| | | | Gaps | 0; |
| Qy | 1 | ATGACAGAGCAAAAGAGGAAAGAAAGCAATGCTTTGGAGCAAGATCATACCCCCCAG | 60 | |
| Db | 1 | ATGACAGAGCAAAAGAGGAAAGAAAGCAATGCTTTGGAGCAAGATCATACCCCCCAG | 60 | |
| Qy | 61 | TGTCAAGACATTATACAGTGTGTAGAGGGAGCTGCCCATTTTGGATTTGGCACCACAT | 120 | |
| Db | 61 | TGTCAAGACATTATACAGTGTGTAGAGGGAGCTGCCCATTTTGGATTTGGCACCACAT | 120 | |
| Qy | 121 | TACATCTGAAAGAAAACCTTCTTCCAGACACTGTGTCTGGGTAATATTTGGCAGTTCAA | 180 | |
| Db | 121 | TACATCTGAAAGAAAACCTTCTTCCAGACACTGTGTCTGGGTAATATTTGGCAGTTCAA | 180 | |
| Qy | 181 | CAGTGACCCCAAGATTTGGCCCTTTCTCTCATCTGTGACCCAGTGTGGTTTGA | 240 | |
| Db | 181 | CAGTGACCCCAAGATTTGGCCCTTTCTCTCATCTGTGACCCAGTGTGGTTTGA | 240 | |
| Qy | 241 | TATGGTCTCTGTTTCTTCGCCATAATTTATSCCATATTTGGAATGGGACATCATGTGCC | 300 | |
| Db | 241 | TATGGTCTCTGTTTCTTCGCCATAATTTATSCCATATTTGGAATGGGACATCATGTGCC | 300 | |

```
QY 301 ACAGGCACCTTTGCTTGCATCTCTTAATATCAGCAACGCGTGGAAACGGAATGTCCT 360
Db 301 ACAGGCACCTTTGCTTGCATCTCTTAATATCAGCAACGCGTGGAAACGGAATGTCCT 360
QY 361 CAGAACATGCGAATCTCACACACAGAGTAACACAGCGTGTGGCTTATCCGACTTT 420
Db 361 CAGAACATGCGAATCTCACACACAGAGTAACACAGCGTGTGGCTTATCCGACTTT 420
QY 421 GAAATGCAAGGATCCAGTGTCTGCAGCAGTTTCTCTTCTGGAGAGTGATTCAGGTG 480
Db 421 GAAATGCAAGGATCCAGTGTCTGCAGCAGTTTCTCTTCTGGAGAGTGATTCAGGTG 480
QY 481 GCCATGTTTGTCTGCAACTGGGAGTGCACATTTGTGGTCCACAGAGCCTGTGATCAGC 540
Db 481 GCCATGTTTGTCTGCAACTGGGAGTGCACATTTGTGGTCCACAGAGCCTGTGATCAGC 540
QY 541 GCAATGACAACCTGGGGCTGCCACCCATGTGTGACTTCACAAGTCAAAATATCTCTGGGA 600
Db 541 GCAATGACAACCTGGGGCTGCCACCCATGTGTGACTTCACAAGTCAAAATATCTCTGGGA 600
QY 601 ATGAAATGCCATATATATCCGACCACTTGGATCTTTTATATTTATGTCATATGTTTTT 660
Db 601 ATGAAATGCCATATATATCCGACCACTTGGATCTTTTATATTTATGTCATATGTTTTT 660
QY 661 GAAACATCAAGTCTGTGCGACTGGAAGCATTTGCTTTATCTTCTGAGCATTTGTGTC 720
Db 661 GAAACATCAAGTCTGTGCGACTGGAAGCATTTGCTTTATCTTGTCTGAGCATTTGTGTC 720
QY 721 CTTGTTCTTTTAAAGAGCTGAATGAACAGTTTAAAGGAAAATTAAGATTGTTCTTCT 780
Db 721 CTTGTTCTTTTAAAGAGCTGAATGAACAGTTTAAAGGAAAATTAAGATTGTTCTTCT 780
QY 781 GTAGATTTAGTTTGTGATTTATGTCATCATTTGCTTATTCGACCAATATGGAAC 840
Db 781 GTAGATTTAGTTTGTGATTTATGTCATCATTTGCTTATTCGACCAATATGGAAC 840
QY 841 ACATATGGATTAGAATAGTTGTCATATTCACAAGGAAATTCCTCACCTAGAGCTCCC 900
Db 841 ACATATGGATTAGAATAGTTGTCATATTCACAAGGAAATTCCTCACCTAGAGCTCCC 900
QY 901 CCGATGAACATCCTCTCTGCGGTGATCATAGAGCTTTGAGATGGCACTTTGTAGGCTAT 960
Db 901 CCGATGAACATCCTCTCTGCGGTGATCATAGAGCTTTGAGATGGCACTTTGTAGGCTAT 960
QY 961 GTGGCTCAGCTGGCTTTGCTCAAGGATCTGCCAAAATTCAAATATTCATTTGATGAC 1020
Db 961 GTGGCTCAGCTGGCTTTGCTCAAGGATCTGCCAAAATTCAAATATTCATTTGATGAC 1020
QY 1021 AACCAAGGAATTTTGGCCCATGGCTCAGCAATATAGTTTCTTCATTTTCTCTGCATA 1080
Db 1021 AACCAAGGAATTTTGGCCCATGGCTCAGCAATATAGTTTCTTCATTTTCTCTGCATA 1080
QY 1081 CCAAGTGTCTGCCATGGGAAGGACGGCTGCTGTAACAGACAGGAGGAGACACAG 1140
Db 1081 CCAAGTGTCTGCCATGGGAAGGACGGCTGCTGTAACAGACAGGAGGAGACACAG 1140
QY 1141 GTGGCTGTCTAAATCTTGCAATTTTCTGCTTATAGTCATCTATGCAATAGGACCTTTG 1200
Db 1141 GTGGCTGTCTAAATCTTGCAATTTTCTGCTTATAGTCATCTATGCAATAGGACCTTTG 1200
QY 1201 CTTTACTGGCTGCCCATGTGTCTCTGCAAGCAATTTATTTGTGGAGCTGAAGGGAATG 1260
Db 1201 CTTTACTGGCTGCCCATGTGTCTCTGCAAGCAATTTATTTGTGGAGCTGAAGGGAATG 1260
QY 1261 CTAATACAGTCCGAGATTTAAAAAATTTGGAAATGGAATGGAATAAATCGATGGGG 1316
Db 1261 CTAATACAGTCCGAGATTTAAAAAATTTGGAAATGGAATGGAATAAATCGATGGGG 1316
```

RESULT 2

US-10-761-033-11
; Sequence 11, Application US/10761033
; Publication No. US20040138442A1

```
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; TITLE OF INVENTION: Novel Human Transporter Proteins and Polynucleotides Encoding the  
; TITLE OF INVENTION: Same  
; FILE REFERENCE: LEX-0186-USA  
; CURRENT APPLICATION NUMBER: US/10/761.033  
; CURRENT FILING DATE: 2004-01-20  
; PRIOR APPLICATION NUMBER: US/09/875,811  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/210,045  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 1797  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-10-761-033-11
```

| Query Match | 100.0% | Score 1316; | DB 7; | Length 1797; |
|-----------------------|--|---------------|-----------|--------------|
| Best Local Similarity | 100.0% | Pred. No. 0; | | |
| Matches 1316; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| QY 1 | ATGACAGAGCAAAAGAGAGAAAAGAAAGCATGCTTTGGAGCAAGATGCATACCCCCCAG | 60 | | |
| Db 1 | ATGACAGAGCAAAAGAGAGAAAAGAAAGCATGCTTTGGAGCAAGATGCATACCCCCCAG | 60 | | |
| QY 61 | TGTGAAGACATTATACAGTGTGTAGAAAGGCACTGCCCATTTTGGATTGGGCAACAT | 120 | | |
| Db 61 | TGTGAAGACATTATACAGTGTGTAGAAAGGCACTGCCCATTTTGGATTGGGCAACAT | 120 | | |
| QY 121 | TACAATCTGAAAGAAAACCTTCTCCAGACACTGTGTCTGGGTAATGTTGCGAGTTCAA | 180 | | |
| Db 121 | TACAATCTGAAAGAAAACCTTCTCCAGACACTGTGTCTGGGTAATGTTGCGAGTTCAA | 180 | | |
| QY 181 | CAGGTGACCCAAAGGATTTGGCTTTGCTCTCATCTGTGCACCCAGTGTGTTGTTTA | 240 | | |
| Db 181 | CAGGTGACCCAAAGGATTTGGCTTTGCTCTCATCTGTGCACCCAGTGTGTTGTTTA | 240 | | |
| QY 241 | TATGGGTCTGTGTTTCTGCGCATTAATTTATGCCATATTTGGAAATGGGACATCATGTTGCC | 300 | | |
| Db 241 | TATGGGTCTGTGTTTCTGCGCATTAATTTATGCCATATTTGGAAATGGGACATCATGTTGCC | 300 | | |
| QY 301 | ACAGGCACCTTTGCTTGCATCTCTTAATATCAGCAACGCGTGGAAACGGAATGTCCT | 360 | | |
| Db 301 | ACAGGCACCTTTGCTTGCATCTCTTAATATCAGCAACGCGTGGAAACGGAATGTCCT | 360 | | |
| QY 361 | CAGAACATGCAAGATCTCACACACAGAGTAACACAGCGTGTGGCTTATCCGACTTT | 420 | | |
| Db 361 | CAGAACATGCAAGATCTCACACACAGAGTAACACAGCGTGTGGCTTATCCGACTTT | 420 | | |
| QY 421 | GAAATGCAAGGATCCAGTGTGTGAGCAGTTTCTTCTTGGAGAGTGTGATTCAGGTG | 480 | | |
| Db 421 | GAAATGCAAGGATCCAGTGTGTGAGCAGTTTCTTCTTGGAGAGTGTGATTCAGGTG | 480 | | |
| QY 481 | GCCATGTTTGTGCTGCAACTGGGAGTGCACATTTGTGGTCCACAGAGCCTGTGATCAGC | 540 | | |
| Db 481 | GCCATGTTTGTGCTGCAACTGGGAGTGCACATTTGTGGTCCACAGAGCCTGTGATCAGC | 540 | | |
| QY 541 | GCAATGACAACCTGGGGCTGCCACCCATGTGTGACTTCACAAGTCAAAATATCTCTGGGA | 600 | | |
| Db 541 | GCAATGACAACCTGGGGCTGCCACCCATGTGTGACTTCACAAGTCAAAATATCTCTGGGA | 600 | | |
| QY 601 | ATGAAATGCCATATATATCCGACCACTTGGATCTTTTATATTTATGTCATATGTTTTT | 660 | | |
| Db 601 | ATGAAATGCCATATATATCCGACCACTTGGATCTTTTATATTTATGTCATATGTTTTT | 660 | | |
| QY 661 | GAAACATCAAGTCTGTGCGACTGGAAGCATTTGCTTTATCTTCTGAGCATTTGTGTC | 720 | | |
| Db 661 | GAAACATCAAGTCTGTGCGACTGGAAGCATTTGCTTTATCTTGTCTGAGCATTTGTGTC | 720 | | |
| QY 721 | CTTGTCTTTTAAAGAGCTGAATGAACAGTTTAAAGGAAAATTAAGATTGTTCTTCT | 780 | | |

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 22:29:16 ; Search time 5828 Seconds
(without alignments)

10564.825 Million cell updates/sec

Title: US-10-761-033-1-copy_1_1316

Perfect score: 1316

Sequence: 1 atgcaggagcaagaggaa.....gtggataaaatcgattgggg 1316

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1.*

2: gb_est2.*

3: gb_est3.*

4: gb_est4.*

5: gb_est5.*

6: gb_est6.*

7: gb_est7.*

8: gb_est8.*

9: gb_est9.*

10: gb_est10.*

11: gb_est11.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|-------------|--------------------|
| 1 | 1291.2 | 98.1 | 5240 | 4 CR859003 | CR859003 Pongo pyg |
| 2 | 1124.8 | 85.5 | 2650 | 4 BC060784 | BC060784 Homo sapi |
| 3 | 1124.8 | 85.5 | 2650 | 4 BC060872 | BC060872 Homo sapi |
| 4 | 1069.6 | 81.3 | 4393 | 4 AK054327 | AK054327 Mus muscu |
| 5 | 962.4 | 73.1 | 1800 | 10 AY405555 | AY405555 Homo sapi |
| 6 | 959.2 | 72.9 | 1800 | 10 AY405556 | AY405556 Pan trogl |
| 7 | 762.4 | 57.9 | 1800 | 10 AY405557 | AY405557 Mus muscu |
| 8 | 631.4 | 48.0 | 645 | 3 BM726084 | BM726084 UI-E-EUO |
| 9 | 611.2 | 46.4 | 729 | 3 BM682518 | BM682518 UI-E-EUO |
| 10 | 562 | 42.7 | 792 | 7 CR629535 | CR629535 DKFZp469J |
| 11 | 556.2 | 42.3 | 799 | 6 CR745519 | CR745519 UI-M-GVO |
| 12 | 524.4 | 39.8 | 773 | 7 CR764738 | CR764738 DKFZp469M |
| 13 | 518.6 | 39.4 | 744 | 7 CR762687 | CR762687 DKFZp469P |
| 14 | 515 | 39.1 | 864 | 6 CB994523 | CB994523 AGENCOURT |
| 15 | 497.4 | 37.8 | 725 | 7 CR549818 | CR549818 DKFZp469J |
| 16 | 469.8 | 35.7 | 715 | 6 CF744029 | CF744029 UI-M-GVO |
| 17 | 467.8 | 35.5 | 787 | 7 CR769333 | CR769333 DKFZp469P |
| 18 | 447.8 | 34.0 | 689 | 7 CR791972 | CR791972 DKFZp469M |
| 19 | 443 | 33.7 | 656 | 7 CR767955 | CR767955 DKFZp469M |
| 20 | 442 | 33.6 | 656 | 7 CR771034 | CR771034 DKFZp469A |
| 21 | 438.2 | 33.3 | 768 | 6 CB989260 | CB989260 AGENCOURT |
| 22 | 426.2 | 32.4 | 939 | 2 BF143423 | BF143423 601790065 |

| | | | | | | |
|----|-------|------|------|----|-----------|--------------------|
| 23 | 424.2 | 32.2 | 638 | 5 | EX484645 | EX484645 DKFZp468N |
| 24 | 407 | 30.9 | 670 | 2 | BB666404 | BB666404 |
| 25 | 373.2 | 28.4 | 875 | 8 | CX342514 | CX342514 JGI X2T46 |
| 26 | 373 | 28.3 | 827 | 7 | CF993927 | CF993927 AGENCOURT |
| 27 | 362.8 | 27.6 | 928 | 6 | CB993059 | CB993059 AGENCOURT |
| 28 | 358.8 | 27.3 | 568 | 7 | CR774608 | CR774608 DKFZp469I |
| 29 | 352.8 | 26.8 | 562 | 7 | CR771087 | CR771087 DKFZp469K |
| 30 | 347.2 | 26.4 | 593 | 7 | CR557400 | CR557400 DKFZp469I |
| 31 | 345.8 | 26.3 | 555 | 7 | CR751908 | CR751908 DKFZp469O |
| 32 | 336.6 | 25.6 | 533 | 7 | CR554938 | CR554938 DKFZp469N |
| 33 | 336.2 | 25.5 | 548 | 7 | CR753714 | CR753714 DKFZp469J |
| 34 | 331.8 | 25.2 | 578 | 7 | CR753649 | CR753649 DKFZp469H |
| 35 | 319.2 | 24.3 | 658 | 7 | CR790935 | CR790935 DKFZp469J |
| 36 | 300.8 | 22.9 | 524 | 6 | CD803846 | CD803846 UI-M-GVO |
| 37 | 299.4 | 22.8 | 418 | 6 | CD701536 | CD701536 EST18060 |
| 38 | 268 | 20.4 | 418 | 7 | CR554171 | CR554171 DKFZp469K |
| 39 | 226.4 | 17.2 | 433 | 7 | CR751437 | CR751437 DKFZp469J |
| 40 | 222.2 | 16.9 | 2140 | 11 | DQ030554 | DQ030554 Homo sapi |
| 41 | 221 | 16.8 | 364 | 5 | BY770037 | BY770037 BY770037 |
| 42 | 221 | 16.8 | 2624 | 4 | HSM806949 | BX640837 Homo sapi |
| 43 | 220.6 | 16.8 | 2878 | 4 | CR926479 | CR926479 Pongo pyg |
| 44 | 209.4 | 15.9 | 3603 | 4 | AK052641 | AK052641 Mus muscu |
| 45 | 207.6 | 15.8 | 2140 | 11 | DQ030555 | DQ030555 Pan trogl |

ALIGNMENTS

| | | | | | |
|------------|--|-------------|------|--------|-----------------|
| RESULT 1 | CR859003 | 5240 bp | mRNA | linear | HTC 17-APR-2005 |
| LOCUS | Pongo pygmaeus mRNA; cDNA DKFZp469J151 (from clone DKFZp469J151). | | | | |
| DEFINITION | CR859003 | | | | |
| ACCESSION | CR859003.1 | GI:55728919 | | | |
| VERSION | HTC | | | | |
| KEYWORDS | Pongo pygmaeus (orangutan) | | | | |
| SOURCE | Pongo pygmaeus | | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Pongo. | | | | |
| REFERENCE | 1 (bases 1 to 5240) | | | | |
| AUTHORS | Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R., Mewes H.W., Well B., Amid C., Osanger A., Fobo G., Han M. and Wiemann S. | | | | |
| CONSRM | The German cDNA Consortium | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY | | | | |
| COMMENT | Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp469J151) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp469J151 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/. | | | | |
| FEATURES | Location/Qualifiers | | | | |
| source | 1..5240 | | | | |
| | /organism="Pongo pygmaeus" | | | | |
| | /mol_type="mRNA" | | | | |
| | /db_xref="taxon:9606" | | | | |
| | /clone="DKFZp469J151" | | | | |
| | /tissue_type="kidney" | | | | |
| | /clone_lib="469 (synonym: pkid1). Vector pSport1_Sfi; host DH10B; sites SfilA + SfilB" | | | | |
| | /dev_stage="adult" | | | | |
| | /note="solute carrier family 26, member 7 isoform a (Homo sapiens)" | | | | |
| gene | 1..5240 | | | | |
| | /gene="DKFZp469J151" | | | | |
| CDS | join(208..228,228..2177) | | | | |

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 22:22:12 ; Search time 724 Seconds
(without alignments)

12114.276 Million cell updates/sec

Title: US-10-761-033-1_COPY_1_1316

Perfect score: 1316

Sequence: 1 atgacaggagcaagaggaa.....gtggataaaatcgattggggg 1316

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

14: Geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1 | 1316 | 100.0 | 1797 | 6 | ABK12981 |
| 2 | 1316 | 100.0 | 1818 | 6 | ABK12979 |
| 3 | 1316 | 100.0 | 1866 | 6 | ABK12977 |
| 4 | 1316 | 100.0 | 1971 | 6 | ABK12980 |
| 5 | 1316 | 100.0 | 1971 | 10 | ACF79733 |
| 6 | 1316 | 100.0 | 1992 | 6 | ABK12978 |
| 7 | 1316 | 100.0 | 2040 | 6 | ABK12976 |
| 8 | 1316 | 100.0 | 2345 | 10 | ADC35055 |
| 9 | 1316 | 100.0 | 2420 | 6 | ABK12982 |
| 10 | 1314.4 | 99.9 | 2631 | 5 | ADL62785 |
| 11 | 1314.4 | 99.9 | 2886 | 10 | ADC35054 |
| 12 | 1314.4 | 99.9 | 5265 | 6 | ADC35215 |
| 13 | 1169.4 | 88.9 | 5312 | 13 | ACN41251 |
| 14 | 1166.2 | 88.6 | 2848 | 6 | ABK83213 |
| 15 | 1068 | 81.2 | 2289 | 10 | ACF79734 |
| 16 | 514.4 | 39.1 | 1296 | 5 | AS70486 |
| 17 | 224.4 | 17.1 | 2692 | 12 | ADH22591 |
| 18 | 224.4 | 17.1 | 2879 | 9 | ADA10954 |
| 19 | 224.4 | 17.1 | 2879 | 12 | ADG47980 |

| | | | | | | |
|----|-------|------|------|----|----------|--------------------|
| 20 | 224.4 | 17.1 | 2882 | 2 | AAQ71399 | Aaq71399 Down-regu |
| 21 | 224.4 | 17.1 | 2908 | 13 | ACN43500 | Acn43500 Human dia |
| 22 | 224.4 | 17.1 | 2918 | 5 | ADL62992 | Adl62992 Human ova |
| 23 | 217.4 | 16.5 | 2343 | 4 | AAI99656 | Aai99656 Human pen |
| 24 | 217.4 | 16.5 | 2782 | 4 | AAI99681 | Aai99681 Mouse pen |
| 25 | 217.4 | 16.5 | 4930 | 6 | ABK92231 | Abk92231 Prostate |
| 26 | 217.4 | 16.5 | 4930 | 11 | ADN39611 | Adn39611 Cancer/an |
| 27 | 217.4 | 16.5 | 4930 | 11 | ADN95181 | Adn95181 Human BSC |
| 28 | 217.4 | 16.5 | 4930 | 12 | ADJ74750 | Adj74750 Human pen |
| 29 | 217.4 | 16.5 | 4930 | 12 | ADJ74895 | Adj74895 Marker ge |
| 30 | 214.2 | 16.3 | 2829 | 11 | ADM56356 | Adm56356 Frog SLC2 |
| 31 | 206.8 | 15.7 | 1835 | 8 | ADA47351 | Ada47351 Human tra |
| 32 | 206.8 | 15.7 | 2378 | 13 | ACN43826 | Acn43826 Human dia |
| 33 | 204 | 15.5 | 3049 | 10 | ACF79737 | Acf79737 Mouse ani |
| 34 | 203.2 | 15.4 | 3635 | 11 | ADM56348 | Adm56348 Frog SLC2 |
| 35 | 201 | 15.3 | 249 | 5 | ADL36859 | Adl36859 Human ova |
| 36 | 201 | 15.3 | 249 | 5 | ADI71704 | Adi71704 Human ova |
| 37 | 201 | 15.3 | 274 | 5 | ADL36868 | Adl36868 Human ova |
| 38 | 201 | 15.3 | 274 | 5 | ADI71713 | Adi71713 Human ova |
| 39 | 201 | 15.3 | 373 | 5 | ADL43243 | Adl43243 Human ova |
| 40 | 201 | 15.3 | 413 | 5 | ADL43252 | Adl43252 Human ova |
| 41 | 199.4 | 15.2 | 2264 | 10 | ADJ95139 | Adj95139 Novel NOV |
| 42 | 199.4 | 15.2 | 2264 | 11 | ADM29327 | Adm29327 Human nov |
| 43 | 199 | 15.1 | 2869 | 13 | ACN43501 | Acn43501 Human dia |
| 44 | 189.2 | 14.4 | 2441 | 11 | ADM57788 | Adm57788 Murine pr |
| 45 | 189.2 | 14.4 | 2441 | 12 | ADB86099 | Ade86099 cDNA enco |

ALIGNMENTS

RESULT 1

ABK12981

ID ABK12981 standard; cDNA; 1797 BP.

AC ABK12981;

DT 23-APR-2002 (first entry)

DB cDNA sequence of the gene encoding novel human protein (NHP) #6.

XX Novel human protein; NHP; transporter protein; polymorphism; gene;

XX mental disorder; biological disease; medical disorder; ss.

OS Homo sapiens.

PH Key Location/Qualifiers

FT CDS 1..1797

FT /*tag= a

FT /product= "Novel human protein (NHP) #6"

PN WO200194583-A2.

XX 13-DEC-2001.

XX 06-JUN-2001; 2001WO-US018393.

XX 07-JUN-2000; 2000US-0210045P.

XX (LEXI-) LEXICON GENETICS INC.

XX Walke DW, Scoville J;

XX WPI; 2002-147673/19.

XX P-PSDB; AAU74917.

XX New human polynucleotides encoding proteins that share sequence
PT similarity with mammalian transporter proteins, useful for determining
PT genomic structures, identifying polymorphisms, or as reagents for
PT diagnosis or drug screening.

XX Disclosure; Page 45-46; 48pp; English.

PS

CC The present invention relates to new isolated nucleic acid molecules
CC comprising a nucleotide sequence encoding a protein that shares sequence
CC similarity with mammalian transporter proteins. The invention also
CC relates to a nucleotide sequence that hybridizes under stringent
CC conditions to the nucleotide sequence comprising 1971 bp (ABK12980) fully
CC defined in the specification or its complement. The isolated nucleic acid
CC and the protein it encodes are useful for identifying a coding sequence
CC and mapping a unique gene to a particular chromosome, identifying and
CC characterizing the temporal and tissue specific expression of a gene,
CC screening a human genomic library, determining the genomic structure of a
CC given locus/allele or designing diagnostic tests. The nucleic acids and
CC proteins are particularly useful for identifying polymorphisms and in
CC amplification assays to detect mutations within the exons, introns and
CC splice sites that can be used in diagnostics and pharmacogenomics. These
CC are also useful for generating antibodies, as reagents in diagnostic
CC assays or for identifying other cellular gene products related to novel
CC human proteins. The nucleotides can be employed as reagents in assays for
CC screening for compounds that can be used as pharmaceutical reagents
CC useful in the therapeutic treatment of mental, biological or medical
CC disorders and diseases. The present nucleic acid sequence encodes novel
CC human protein #6 that is one of several (AU74912- AU74917) novel human
CC proteins (NHP) of the invention
XX

50 Sequence 1797 BP; 500 A; 367 C; 396 G; 534 T; 0 U; 0 Other;

Query Match 100.0%; Score 1316; DB 6; Length 1797;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| Qy | 1 | ATGACAGGAGCAAGAGGAAAGAAAGAAAGCAATGCTTTGGAGCAAGATGATACACCCCCAG | 60 |
| Db | 1 | ATGACAGGAGCAAGAGGAAAGAAAGAAAGCAATGCTTTGGAGCAAGATGATACACCCCCAG | 60 |
| Qy | 61 | TGTGAAGACATATACAGTGTGTAGAAAGGCGATGCCCCATTTGGATTGGGCGACCAT | 120 |
| Db | 61 | TGTGAAGACATATACAGTGTGTAGAAAGGCGATGCCCCATTTGGATTGGGCGACCAT | 120 |
| Qy | 121 | TACAACTCTGAAGAAAGAACTTGTCTCCAGACACTGTGTCTGGGTAATGCTGCGAGTTCAA | 180 |
| Db | 121 | TACAACTCTGAAGAAAGAACTTGTCTCCAGACACTGTGTCTGGGTAATGCTGCGAGTTCAA | 180 |
| Qy | 181 | CAGGTGACCCCAAGGATTTGGCCCTTTGCTGTTCTCTCATCTGTGCAACCCAGTGTGGTTTA | 240 |
| Db | 181 | CAGGTGACCCCAAGGATTTGGCCCTTTGCTGTTCTCTCATCTGTGCAACCCAGTGTGGTTTA | 240 |
| Qy | 241 | TATGGGTCTCTGTTTCTGCTGCCATTAATTTATGCCATATTTGGAAATGGGACATCATGTTGCC | 300 |
| Db | 241 | TATGGGTCTCTGTTTCTGCTGCCATTAATTTATGCCATATTTGGAAATGGGACATCATGTTGCC | 300 |
| Qy | 301 | ACAGGCACCTTTGCTTGCATCTCTTAATATCAGCCACGCGGTGGAAACGGATTTGTCCT | 360 |
| Db | 301 | ACAGGCACCTTTGCTTGCATCTCTTAATATCAGCCACGCGGTGGAAACGGATTTGTCCT | 360 |
| Qy | 361 | CAGAACATGCAAGATCTCACACACAGAGTAACACAGCGTGTGGGCTTATCCGACTTT | 420 |
| Db | 361 | CAGAACATGCAAGATCTCACACACAGAGTAACACAGCGTGTGGGCTTATCCGACTTT | 420 |
| Qy | 421 | GAATGCAAGGATCCAGTGTGTCAGCAGTTCCTTCTTGGGAGGTGTGATTCAGGTG | 480 |
| Db | 421 | GAATGCAAGGATCCAGTGTGTCAGCAGTTCCTTCTTGGGAGGTGTGATTCAGGTG | 480 |
| Qy | 481 | GCCATGTTTGTGTCGAACCTGGGAGTGCACATTTTGGTTCACAGAGCCTGTGATCAGC | 540 |
| Db | 481 | GCCATGTTTGTGTCGAACCTGGGAGTGCACATTTTGGTTCACAGAGCCTGTGATCAGC | 540 |
| Qy | 541 | GCAATGACAACTGGGGCTGCCACCCATGTGTGACTTTCACAAGTCAAAATATCTCTTGGGA | 600 |
| Db | 541 | GCAATGACAACTGGGGCTGCCACCCATGTGTGACTTTCACAAGTCAAAATATCTCTTGGGA | 600 |
| Qy | 601 | ATGAATGCAATATATATCCGACCACTTGGATCTTTTATATTTATGATATGTTT | 660 |
| Db | 601 | ATGAATGCAATATATATCCGACCACTTGGATCTTTTATATTTATGATATGTTT | 660 |

| | | | |
|----|------|--|------|
| Qy | 661 | GAAGAACATCAAGTCTGTGCGAGCTGGAAGCATTTCTTTATCTTGTGTCGAGCATTTGTGTC | 720 |
| Db | 661 | GAAGAACATCAAGTCTGTGCGAGCTGGAAGCATTTCTTTATCTTGTGTCGAGCATTTGTGTC | 720 |
| Qy | 721 | CTTGTCTTTGTTAAAGAGCTGAATGAACAGTTTAAAGGAAAAATTAAGTTGTTCTTCT | 780 |
| Db | 721 | CTTGTCTTTGTTAAAGAGCTGAATGAACAGTTTAAAGGAAAAATTAAGTTGTTCTTCT | 780 |
| Qy | 781 | GTAGATTTAGTTTGTATTTGTCGATCATTTGCTGTTTATTCGACCAATATGGAAGAAC | 840 |
| Db | 781 | GTAGATTTAGTTTGTATTTGTCGATCATTTGCTGTTTATTCGACCAATATGGAAGAAC | 840 |
| Qy | 841 | ACATATCGATTTAGATTTGTTGGTTCATATTCACAGGAATTCCTTCACCTAGAGCTCCC | 900 |
| Db | 841 | ACATATCGATTTAGATTTGTTGGTTCATATTCACAGGAATTCCTTCACCTAGAGCTCCC | 900 |
| Qy | 901 | CCGATGAACATCTCTCTGCGGTGATCACTGAAGCTTTCCGAGTGGCACTTTGTAGGCTAT | 960 |
| Db | 901 | CCGATGAACATCTCTCTGCGGTGATCACTGAAGCTTTCCGAGTGGCACTTTGTAGGCTAT | 960 |
| Qy | 961 | GTGGCCTCAGTGGCTCTTGTCTCAAGGATCTGCCAAAAAATTCAAATATTCATTTGATGAC | 1020 |
| Db | 961 | GTGGCCTCAGTGGCTCTTGTCTCAAGGATCTGCCAAAAAATTCAAATATTCATTTGATGAC | 1020 |
| Qy | 1021 | AACACAGAAATTTTGGCCCATGGCTCAGCAATATAGTTTCTTCAATTTTCTTCTGCAATA | 1080 |
| Db | 1021 | AACACAGAAATTTTGGCCCATGGCTCAGCAATATAGTTTCTTCAATTTTCTTCTGCAATA | 1080 |
| Qy | 1081 | CCAAGTCTCTGTCATATCTTGCATTTTCTCTTATAGTCATCTATGCAATAGGACCTTTG | 1140 |
| Db | 1081 | CCAAGTCTCTGTCATATCTTGCATTTTCTCTTATAGTCATCTATGCAATAGGACCTTTG | 1140 |
| Qy | 1141 | GTGGCTTGTCTAATATCTTGCATTTTCTCTTATAGTCATCTATGCAATAGGACCTTTG | 1200 |
| Db | 1141 | GTGGCTTGTCTAATATCTTGCATTTTCTCTTATAGTCATCTATGCAATAGGACCTTTG | 1200 |
| Qy | 1201 | CTTTACTGGCTGCCCATGTGTCTTTCAGCAATATTTGTGTGGGACTGGAAGGAATG | 1260 |
| Db | 1201 | CTTTACTGGCTGCCCATGTGTCTTTCAGCAATATTTGTGTGGGACTGGAAGGAATG | 1260 |
| Qy | 1261 | CTAATACAGTTCGAGATTTTAAAAAATATTTGGAATGTGGAATAAATCGATTTGGG | 1316 |
| Db | 1261 | CTAATACAGTTCGAGATTTTAAAAAATATTTGGAATGTGGAATAAATCGATTTGGG | 1316 |

RESULT 2

ABK12979
ID ABK12979 standard; cDNA; 1818 BP.

XX ABK12979;

XX 23-APR-2002 (first entry)

XX cDNA sequence of the gene encoding novel human protein (NHP) #4.

XX Novel human protein; NHP; transporter protein; polymorphism; gene;
XX mental disorder; biological disease; medical disorder; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1818

FT /*tag= a

FT /product= "Novel human protein (NHP) #4"

XX W0200194583-A2.

XX 13-DEC-2001.

XX 06-JUN-2001; 2001WO-US018393.

XX 07-JUN-2000; 2000US-0210045P.

XX

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 22:24:21 ; Search time 6918 Seconds
(without alignments)

10813.238 Million cell updates/sec

Title: US-10-761-033-1_COPY_1_1316

Perfect score: 1316

Sequence: 1 atgacaggagcaagaggaa.....gtggataaaatcgattgggg 1316

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.in.*

3: gb.env.*

4: gb.on.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pr.*

9: gb.ro.*

10: gb.sts.*

11: gb.sy.*

12: gb.un.*

13: gb.vi.*

14: gb.htg.*

15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 1316 | 100.0 | 1797 | 6 | AR486412 Sequence |
| 2 | 1316 | 100.0 | 1797 | 6 | AX338603 Sequence |
| 3 | 1316 | 100.0 | 1818 | 6 | AR486410 Sequence |
| 4 | 1316 | 100.0 | 1818 | 6 | AX338599 Sequence |
| 5 | 1316 | 100.0 | 1866 | 6 | AR486408 Sequence |
| 6 | 1316 | 100.0 | 1866 | 6 | AX338595 Sequence |
| 7 | 1316 | 100.0 | 1971 | 6 | AR486411 Sequence |
| 8 | 1316 | 100.0 | 1971 | 6 | AX338601 Sequence |
| 9 | 1316 | 100.0 | 1992 | 6 | AR486409 Sequence |
| 10 | 1316 | 100.0 | 1992 | 6 | AX338597 Sequence |
| 11 | 1316 | 100.0 | 2040 | 6 | AR486407 Sequence |
| 12 | 1316 | 100.0 | 2040 | 6 | AX338593 Sequence |
| 13 | 1316 | 100.0 | 2345 | 6 | AX960423 Sequence |
| 14 | 1316 | 100.0 | 2345 | 8 | HSA413230 Homo sapi |
| 15 | 1316 | 100.0 | 2420 | 6 | AR486413 Sequence |
| 16 | 1316 | 100.0 | 2420 | 6 | AX338605 Sequence |
| 17 | 1316 | 100.0 | 2928 | 8 | BC094730 Homo sapi |
| 18 | 1316 | 100.0 | 5254 | 8 | HSA413228 Homo sapi |

| | | | | | | |
|----|--------|------|------|----|-----------|--------------------|
| 19 | 1314.4 | 99.9 | 1971 | 6 | AX465644 | AX465644 Sequence |
| 20 | 1314.4 | 99.9 | 2631 | 6 | CQ413926 | CQ413926 Sequence |
| 21 | 1314.4 | 99.9 | 2886 | 6 | AX960422 | AX960422 Sequence |
| 22 | 1314.4 | 99.9 | 2896 | 8 | HSA413229 | AJ413229 Homo sapi |
| 23 | 1314.4 | 99.9 | 5250 | 8 | AF331521 | AF331521 Homo sapi |
| 24 | 1314.4 | 99.9 | 5267 | 6 | AX465642 | AX465642 Sequence |
| 25 | 1166.2 | 88.6 | 2848 | 6 | AX478094 | AX478094 Sequence |
| 26 | 1068 | 81.2 | 2289 | 9 | AF345194 | AF345194 Mus muscu |
| 27 | 1068 | 81.2 | 2366 | 9 | BC026928 | BC026928 Mus muscu |
| 28 | 627.8 | 47.7 | 1935 | 6 | CQ739120 | CQ739120 Sequence |
| 29 | 328.2 | 24.9 | 385 | 4 | AX166770 | AX166770 Oryctolag |
| 30 | 288.4 | 21.9 | 4143 | 8 | AX122933 | AX122933 Homo sapi |
| 31 | 236.2 | 17.9 | 3065 | 5 | AB084425 | AB084425 Anguilla |
| 32 | 224.4 | 17.1 | 2855 | 6 | CQ714572 | CQ714572 Sequence |
| 33 | 224.4 | 17.1 | 2881 | 8 | HUMDRA | L02785 Homo sapien |
| 34 | 224.4 | 17.1 | 2881 | 10 | G18336 | G18336 sWS1883 Er |
| 35 | 224.4 | 17.1 | 2882 | 6 | AR052312 | AR052312 Sequence |
| 36 | 224.4 | 17.1 | 2882 | 6 | AR144265 | AR144265 Sequence |
| 37 | 224.4 | 17.1 | 2882 | 6 | I28252 | I28252 Sequence 1 |
| 38 | 224.4 | 17.1 | 2918 | 6 | CQ414133 | CQ414133 Sequence |
| 39 | 224.4 | 17.1 | 2955 | 8 | BC025671 | BC025671 Homo sapi |
| 40 | 224.2 | 17.0 | 2629 | 9 | AF136751 | AF136751 Mus muscu |
| 41 | 219.8 | 16.7 | 2904 | 4 | AF314819 | AF314819 Oryctolag |
| 42 | 217.4 | 16.5 | 2343 | 6 | BD017205 | BD017205 Utilizati |
| 43 | 217.4 | 16.5 | 2782 | 6 | BD017230 | BD017230 Utilizati |
| 44 | 217.4 | 16.5 | 4930 | 6 | CQ776316 | CQ776316 Sequence |
| 45 | 217.4 | 16.5 | 4930 | 6 | CQ776461 | CQ776461 Sequence |

ALIGNMENTS

RESULT 1
AR486412
LOCUS AR486412 1797 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 11 from patent US 6703495.
ACCESSION AR486412
VERSION AR486412.1 GI:47251211
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1797)
AUTHORS Walke,D.W. and Scoville,J.
TITLE Polynucleotides encoding human transporter protein
JOURNAL Patent: US 6703495-A 11 09-MAR-2004;
Lexicon Genetics Incorporated; The Woodlands, TX;
WOX;

FEATURES
source Location/Qualifiers
1..1797
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

| | | | | |
|-----------------------|--------------|---|------------|--------------|
| Query Match | 100.0%; | Score 1316; | DB 6; | Length 1797; |
| Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| Matches 1316; | Conservative | 0; | Mismatches | 0; |
| Indels | 0; | Gaps | 0; | |
| Qy | 1 | ATGACAGGCGCAAGCGAAGAAAGAAAGCAATGCTTTGGAGCAAGATGCATACCCCCAG | 60 | |
| Db | 1 | ATGACAGGCGCAAGCGAAGAAAGAAAGCAATGCTTTGGAGCAAGATGCATACCCCCAG | 60 | |
| Qy | 61 | TGTGAAGACATTTATACAGTGGTGTAGAGGGGACATGCCCCATTTTGGATTTGGGCACCAT | 120 | |
| Db | 61 | TGTGAAGACATTTATACAGTGGTGTAGAGGGGACATGCCCCATTTTGGATTTGGGCACCAT | 120 | |
| Qy | 121 | TACAATCTGAAGAAAACTTCTTCCAGACATGTGTCTGGGATAATTTGGCAGTTCAA | 180 | |
| Db | 121 | TACAATCTGAAGAAAACTTCTTCCAGACATGTGTCTGGGATAATTTGGCAGTTCAA | 180 | |
| Qy | 181 | CAGTGACCCCAAGGATTTGGCCCTTTGCTTCTCTCATCTGTGACCCCAAGTTTGGTTTA | 240 | |
| Db | 181 | CAGTGACCCCAAGGATTTGGCCCTTTGCTTCTCTCATCTGTGACCCCAAGTTTGGTTTA | 240 | |

| | | | |
|----|------|---|------|
| Qy | 241 | TATGGGCTCTGTGTTCTCGCCCAATAATTTATGCCATATATTTGGAAATGGGACATCATGTTGCC | 300 |
| Db | 241 | TATGGGCTCTGTGTTCTCGCCCAATAATTTATGCCATATATTTGGAAATGGGACATCATGTTGCC | 300 |
| Qy | 301 | ACAGGCACCTTTGGCTTGACATCCTTTAATATCAGCCAAACGCGTGGAAACGGATATGTCCTCT | 360 |
| Db | 301 | ACAGGCACCTTTGGCTTGACATCCTTTAATATCAGCCAAACGCGTGGAAACGGATATGTCCTCT | 360 |
| Qy | 361 | CAGAACATGCAAGATCTCCACACACAGAGTAGCAACACGCTGCTGGGCTTATCCGACTTT | 420 |
| Db | 361 | CAGAACATGCAAGATCTCCACACACAGAGTAGCAACACGCTGCTGGGCTTATCCGACTTT | 420 |
| Qy | 421 | GAATGCAAGAGATCCAGTTGCTGACAGCTTTCTTTTGGGAGGATGATTCAGGTG | 480 |
| Db | 421 | GAATGCAAGAGATCCAGTTGCTGACAGCTTTCTTTTGGGAGGATGATTCAGGTG | 480 |
| Qy | 481 | GCCATGTTGTGCTGCAACTGGGCACTGCCATTTGTGGTGCACAGAGCCTGTGATCAGC | 540 |
| Db | 481 | GCCATGTTGTGCTGCAACTGGGCACTGCCATTTGTGGTGCACAGAGCCTGTGATCAGC | 540 |
| Qy | 541 | GCAATGCAACTGGGGCTGCCACCATGTGTGTGACTTCACAAGTCAAAATATCTCTTGGBA | 600 |
| Db | 541 | GCAATGCAACTGGGGCTGCCACCATGTGTGTGACTTCACAAGTCAAAATATCTCTTGGBA | 600 |
| Qy | 601 | ATGAAATGGCCATATATATCCGGACCACTTGGATTTCTTTTATATTTATGCAATATGTTTTT | 660 |
| Db | 601 | ATGAAATGGCCATATATATCCGGACCACTTGGATTTCTTTTATATTTATGCAATATGTTTTT | 660 |
| Qy | 661 | GAAGACATCAAGTCTGTGCGACTGGAAGCATTTGCTTTTATCCTTGTGAGCATTTGTGTC | 720 |
| Db | 661 | GAAGACATCAAGTCTGTGCGACTGGAAGCATTTGCTTTTATCCTTGTGAGCATTTGTGTC | 720 |
| Qy | 721 | CTTCTCTTGTAAAGACTGAAATGACAGTTTAAAGGAAAATTTAAAGTTGTTCTTCTCT | 780 |
| Db | 721 | CTTCTCTTGTAAAGACTGAAATGACAGTTTAAAGGAAAATTTAAAGTTGTTCTTCTCT | 780 |
| Qy | 781 | GTAGATTTAGTTTTGATTTATTTGCTGCACTATTGCTTTATTTGCAACATATGGAANAAC | 840 |
| Db | 781 | GTAGATTTAGTTTTGATTTATTTGCTGCACTATTGCTTTATTTGCAACATATGGAANAAC | 840 |
| Qy | 841 | ACATATGGAATTAGAAGTAGTTGGTCAATATCCACAGGAATTCCTCTACCTAGAGCTCCC | 900 |
| Db | 841 | ACATATGGAATTAGAAGTAGTTGGTCAATATCCACAGGAATTCCTCTACCTAGAGCTCCC | 900 |
| Qy | 901 | CCGATGAACATCTCTCTCGCGTGATCACTGAAGCTTTCCGAGTGGCACTTTGAGGCTAT | 960 |
| Db | 901 | CCGATGAACATCTCTCTCGCGTGATCACTGAAGCTTTCCGAGTGGCACTTTGAGGCTAT | 960 |
| Qy | 961 | GTGGCTCTCACTGGCTCTTGCTCAAGGATCTGCCAAAAAATTTCAAATTTGATGAC | 1020 |
| Db | 961 | GTGGCTCTCACTGGCTCTTGCTCAAGGATCTGCCAAAAAATTTCAAATTTGATGAC | 1020 |
| Qy | 1021 | AACCAGGAATTTTTGGCCCCATGGGCTCAGCAATATAGTTTCTCAATTTTCTCTGCAATA | 1080 |
| Db | 1021 | AACCAGGAATTTTTGGCCCCATGGGCTCAGCAATATAGTTTCTCAATTTTCTCTGCAATA | 1080 |
| Qy | 1081 | CCAAAGTCTGCTGCCATGGGAAGGACGGCTGGCTGTACAGCACAGGACGGAACACACAG | 1140 |
| Db | 1081 | CCAAAGTCTGCTGCCATGGGAAGGACGGCTGGCTGTACAGCACAGGACGGAACACACAG | 1140 |
| Qy | 1141 | GTGGCTTGCTTAATATCTTTCATTTTGTCTTATAGTCATCTATGCAATAGGACCTTTG | 1200 |
| Db | 1141 | GTGGCTTGCTTAAATATCTTTCATTTTGTCTTATAGTCATCTATGCAATAGGACCTTTG | 1200 |
| Qy | 1201 | CTTTTACTGGGCTGCCCATGTGTGCTCTGCAAGCATTTATTTGTGGGACTCAAGGGAATG | 1260 |
| Db | 1201 | CTTTTACTGGGCTGCCCATGTGTGCTCTGCAAGCATTTATTTGTGGGACTCAAGGGAATG | 1260 |
| Qy | 1261 | CTAATACAGTTCCGAGATTTTAAAAAATAATTTGGAAATGGGATAAAATCGATTTGGG | 1316 |
| Db | 1261 | CTAATACAGTTCCGAGATTTTAAAAAATAATTTGGAAATGGGATAAAATCGATTTGGG | 1316 |

| RESULT 2 | AX338603 | 1797 bp | DNA | linear | PAT 09-JAN-2002 |
|------------|---|---|---------------|-----------|-----------------|
| LOCUS | AX338603 | Sequence 11 from Patent WO0194583. | | | |
| DEFINITION | AX338603 | GI:18128993 | | | |
| ACCESSION | AX338603.1 | | | | |
| VERSION | 1 | | | | |
| KEYWORDS | Homo sapiens (human) | | | | |
| SOURCE | Homo sapiens | | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo. | | | | |
| REFERENCE | 1 | Walker, D.W. and Scoville, J. | | | |
| AUTHORS | Novel human transporter proteins and polynucleotides encoding the same | | | | |
| TITLE | Patent: WO 0194583-A 11 13-DEC-2001; | | | | |
| JOURNAL | Lexicon Genetics Incorporated (US) | | | | |
| FEATURES | 1. 1797 | | | | |
| source | /organism="Homo sapiens" | | | | |
| | /mol_type="unassigned DNA" | | | | |
| | /db_xref="taxon:9606" | | | | |
| ORIGIN | | | | | |
| | Query Match | 100.0%; | Score 1316; | DB 6; | Length 1797; |
| | Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| | Matches 1316; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 1 | ATGCAGGCGCAAGAGGAAAGCAATGCTTTGGAGCAAGATGCATACCCCCAG | 60 | | |
| Db | 1 | ATGCAGGCGCAAGAGGAAAGCAATGCTTTGGAGCAAGATGCATACCCCCAG | 60 | | |
| Qy | 61 | TGTGAAGACATTATACAGTGGTGTAGAAGCGGACTGCCATTTTGGATTGGGCAACACAT | 120 | | |
| Db | 61 | TGTGAAGACATTATACAGTGGTGTAGAAGCGGACTGCCATTTTGGATTGGGCAACACAT | 120 | | |
| Qy | 121 | TACAATCTGAAGAAACCTTGCTTCGAGACACATGTGTCTGGGATATATGTTGGCAGTTCAA | 180 | | |
| Db | 121 | TACAATCTGAAGAAACCTTGCTTCGAGACACATGTGTCTGGGATATATGTTGGCAGTTCAA | 180 | | |
| Qy | 181 | CAGGTGACCAAGATGGGCTTTGCTGTCTCATCTGTGCACCCAGTGTGTTGTTTA | 240 | | |
| Db | 181 | CAGGTGACCAAGATGGGCTTTGCTGTCTCATCTGTGCACCCAGTGTGTTGTTTA | 240 | | |
| Qy | 241 | TATGGGTCTCTGTTTCTCGCCATAATTTATGCCATATTTGGATGGGACATCATGTTGCC | 300 | | |
| Db | 241 | TATGGGTCTCTGTTTCTCGCCATAATTTATGCCATATTTGGATGGGACATCATGTTGCC | 300 | | |
| Qy | 301 | ACAGGCACCTTTGGCTTGACATCTTAATATCAGCCAAAGCGGTGGAAAGGATGTCCT | 360 | | |
| Db | 301 | ACAGGCACCTTTGGCTTGACATCTTAATATCAGCCAAAGCGGTGGAAAGGATGTCCT | 360 | | |
| Qy | 361 | CAGAAATGAGAAATCTCACACACAGAGTAAACACAGCGTGTGGGCTTATCCGACTTT | 420 | | |
| Db | 361 | CAGAAATGAGAAATCTCACACACAGAGTAAACACAGCGTGTGGGCTTATCCGACTTT | 420 | | |
| Qy | 421 | GAATGCAAAAGGATCCACGTTGTGTGCAGCAGTTTCTCTTTGGGAGGTGTGATTCAGGTG | 480 | | |
| Db | 421 | GAATGCAAAAGGATCCACGTTGTGTGCAGCAGTTTCTCTTTGGGAGGTGTGATTCAGGTG | 480 | | |
| Qy | 481 | GCCATGTTGTGTCGAACTGGGAGTGCCACATTTGTGTGTACAGAGCTGTGATCAGC | 540 | | |
| Db | 481 | GCCATGTTGTGTCGAACTGGGAGTGCCACATTTGTGTGTACAGAGCTGTGATCAGC | 540 | | |
| Qy | 541 | GCAATGCAACTGGGGTGCACCCATGTGGTCACTTCAACAGTCAAATATCTTTGGGA | 600 | | |
| Db | 541 | GCAATGCAACTGGGGTGCACCCATGTGGTCACTTCAACAGTCAAATATCTTTGGGA | 600 | | |
| Qy | 601 | ATGAAATGCCATATATATCCGACCACTTGGATTCTTTTATATTTATGCATATGTTTT | 660 | | |
| Db | 601 | ATGAAATGCCATATATATCCGACCACTTGGATTCTTTTATATTTATGCATATGTTTT | 660 | | |

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2006, 18:49:38 ; Search time 24 Seconds
(without alignments)
509.146 Million cell updates/sec

Title: US-10-761-033-2_COPY_1_439

Perfect score: 2332
Sequence: 1 MTGAKRKKKMLMSKMTTPQ.....MLIQFDRLKKYNNVDKIDWG 439

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 161667 seqs, 27834885 residues

Total number of hits satisfying chosen parameters: 161667

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA New:*
- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--|
| 1 | 805 | 36.1 | 735 | 6 | US-10-505-263-89 Sequence 89, Appl |
| 2 | 774 | 34.7 | 739 | 6 | US-10-505-263-12 Sequence 12, Appl |
| 3 | 753.5 | 33.8 | 739 | 6 | US-10-505-263-96 Sequence 96, Appl |
| 4 | 742.5 | 33.3 | 719 | 6 | US-10-505-263-87 Sequence 87, Appl |
| 5 | 734 | 32.9 | 753 | 6 | US-10-505-263-91 Sequence 91, Appl |
| 6 | 732 | 32.8 | 738 | 6 | US-10-505-263-4 Sequence 4, Appl |
| 7 | 732 | 32.8 | 758 | 6 | US-10-505-263-2 Sequence 2, Appl |
| 8 | 725.5 | 32.5 | 735 | 6 | US-10-505-263-8 Sequence 8, Appl |
| 9 | 725.5 | 32.5 | 758 | 6 | US-10-505-263-6 Sequence 6, Appl |
| 10 | 691.5 | 31.0 | 778 | 6 | US-10-505-263-81 Sequence 81, Appl |
| 11 | 686.5 | 30.8 | 769 | 6 | US-10-505-263-83 Sequence 83, Appl |
| 12 | 685 | 30.7 | 788 | 6 | US-10-505-263-85 Sequence 85, Appl |
| 13 | 654 | 29.3 | 704 | 6 | US-10-505-263-10 Sequence 10, Appl |
| 14 | 645.5 | 28.9 | 791 | 6 | US-10-537-002-65 Sequence 65, Appl |
| 15 | 431 | 19.3 | 616 | 7 | US-11-096-568A-33957 Sequence 33957, A |
| 16 | 431 | 19.3 | 685 | 7 | US-11-096-568A-33956 Sequence 33956, A |
| 17 | 431 | 19.3 | 717 | 7 | US-11-096-568A-33955 Sequence 33955, A |
| 18 | 418 | 18.7 | 93 | 6 | US-10-505-263-76 Sequence 76, Appl |
| 19 | 396 | 17.7 | 634 | 7 | US-11-096-568A-31697 Sequence 31697, A |
| 20 | 396 | 17.7 | 669 | 7 | US-11-096-568A-31696 Sequence 31696, A |
| 21 | 392 | 17.6 | 658 | 7 | US-11-096-568A-703 Sequence 703, Appl |
| 22 | 369.5 | 16.6 | 677 | 7 | US-11-010-239-48 Sequence 48, Appl |
| 23 | 366.5 | 16.4 | 560 | 7 | US-11-096-568A-31698 Sequence 31698, A |
| 24 | 357 | 16.0 | 649 | 7 | US-11-010-239-44 Sequence 44, Appl |
| 25 | 285.5 | 12.8 | 518 | 7 | US-11-096-568A-704 Sequence 704, Appl |

ALIGNMENTS

RESULT 1
US-10-505-263-89
; Sequence 89, Application US/10505263
; Publication No. US20060014940A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Case Western Reserve University
; APPLICANT: The Brigham and Women's Hospital, Inc.
; APPLICANT: Mount, David B
; APPLICANT: Romero, Michael
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF SLC26A6, SLC26A1, and SLC26A2
; FILE REFERENCE: 1242/50/2 PCT/US
; CURRENT APPLICATION NUMBER: US/10/505,263
; CURRENT FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: US 60/360,275
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US03/06469
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 89
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-10-505-263-89

Query Match 36.1%; Score 805; DB 6; Length 735;
Best Local Similarity 39.2%; Pred. No. 6.1e-56;
Matches 159; Conservative 91; Mismatches 154; Indels 2; Gaps 1;

| | | | |
|----|-----|--|-----|
| Qy | 32 | LPIDWAPHYLNKELLPDVTSGIMLAVQVOTQGLAPAVLSSVHPVFGYGLSLPPIIYA | 91 |
| Db | 67 | IPILGMLPRPVKKEWLLGDIVSLGSLVGIQLPQGLAVALLAGVPVFGYGLSFPFVLIIYA | 126 |
| Qy | 92 | IFGMGHVATGTFTALTSLSANAVRIVPQNMQLTTQSNTSVLGLSDPFEMQRIHVAAV | 151 |
| Db | 127 | IFGTSRHSIPCTFAVISVMVGSVTESLVPS--ENVRLPGNESVIDIARDNDRVEVASAL | 184 |
| Qy | 152 | SFLGGVQVAMFVLQSGATFVVTPEVISAMTGAATHVTSQVKYLLGMKMPYISGPIG | 211 |
| Db | 185 | TFLVGLFQIMLGLVQGVFVVTYLSPLIRGTYSAAAIHVTVSQMKSVLGVQISQKSHPLS | 244 |
| Qy | 212 | PFYIYAVFENIKSVRLKALLLSLISVLVVKELNEQPKRKIKVLPVPLVLIIAASF | 271 |
| Db | 245 | LIYAFVNLCAKVPETNTASLLIGCISITVLPVLFKNDKYSKRMPIELITIVATG | 304 |
| Qy | 272 | ACYCTNNITYGLEVVGVHPIQGISPRAPPMNILLSAVITEAFGVALGVYVASLAAQSGA | 331 |

Db 305 ISYGASLHQVGVGDIVGPIGTGMKAPMLPNTNIPARVVGNAFAIAVAVVYATFISLAKNFG 364
Qy 332 KFKYSIDNQBFLAHLGNISVSSPPFCIPSAAMGRTAGLYSTGAKTQVACLISCIFVL 391
Db 365 VKHGYNDNBQBLIALGSLSIGSPFQCFITGTAMSRSLVQBSTGHSQVASAVSSLVIL 424
Qy 392 IVIYAIGPLLYWLMPCVLASIVVGLKGMILQFRDLKKYWNVDKID 437
Db 425 IILLKAGLPELTPKAILAAVVVNLKGIYQFTDVPMLWRSNKFD 470

RESULT 2
US-10-505-263-12
; Sequence 12, Application US/10505263
; Publication No. US20060014940A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Case Western Reserve University
; APPLICANT: The Brigham and Women's Hospital, Inc.
; APPLICANT: Mount, David B
; APPLICANT: Romero, Michael
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF SLC26A6, SLC26A1, and SLC26A2
; FILE REFERENCE: 1242/50/2 PCT/US
; CURRENT APPLICATION NUMBER: US/10/505,263
; PRIOR FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: US 60/360,275
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US03/06469
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-505-263-12

Query Match 34.7%; Score 774; DB 6; Length 739;
Best Local Similarity 33.1%; Pred. No. 1.7e-53;
Matches 155; Conservative 108; Mismatches 173; Indels 32; Gaps 2;

Qy 2 TGAKRKKSLMSKXHTPQCEDIQWCRRL-----PILDWAPHYN 42
Db 44 TQGERAPRRTHMELREXPDDITQFVIRELQKSCQSAKVRDGPDPFVLRWLPKYD 103
Qy 43 LKENLLPDTVSGIMLAVQVQVTLQGLAFVLSVHPVGLYSLPAIIYVAFGMGHVATG 102
Db 104 LKKNILGDMVMSGLIVGILLVPOSIAYSLLAGOEPYGLYTSFFASIIYFLPFTSRHSIVG 163
Qy 103 TPALTSLISANAVRIVPQNQNLTQS-----NTSVGLSDPEMQRHVA 149
Db 164 IFGLILGIVGVRDLHKACPDTDATSSIAVSSGCVVNNHTLDGLCDKSCYAIKIGS 223
Qy 150 AVSFLGGYIVQAMFVLQGSATFVVTEPVIAMTGAATHVVTQVKLLGMKMPYISGP 209
Db 224 TVTFWAGYVQVAMGPFQGVSVVLSDDLGSFTLTSGAKYLLGLSLPRESHGV 283
Qy 210 LGPFYIYAVPENIKSVRLLEALLSLLSVLVVLKELNEQFKRKIKVLPDVLVLI 269
Db 284 GSVITTWIHIENRINTNICDLITSLCLLVLPSEKELNEHFKDLKAPIPELVIVVAA 343
Qy 270 SPACYCTNMNTYGLVGVGHIPQIPSPRAPMNLISAVITEAGVALGVVASLAAQ 329
Db 344 TLASHFGKLGNGYNSIAGHPTGFMPPKAPDWSLIPNVAIDAISIGFAITVLSLEM 403
Qy 330 SAKFKPKYIDNQBFLAHLGNISVSSPPFCIPSAAMGRTAGLYSTGAKTQVACLISCIF 389
Db 404 FAKKHGTVKANQEMYAIQFCNIIIPSPFHCTTSAALAKTLVKESTGCTQLSAIVTALV 463
Qy 390 VLIYIYATGPLLYLWLMPCVLASIIIVVGLKGMILQFRDLKKYWNVDKID 437

Db 464 LLLVLLVIAPLFYSLQKCVLGVITIVNLRGALLKFRDLPKMWRLSRMD 511

RESULT 3
US-10-505-263-96
; Sequence 96, Application US/10505263
; Publication No. US20060014940A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Case Western Reserve University
; APPLICANT: The Brigham and Women's Hospital, Inc.
; APPLICANT: Mount, David B
; APPLICANT: Romero, Michael
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF SLC26A6, SLC26A1, and SLC26A2
; FILE REFERENCE: 1242/50/2 PCT/US
; CURRENT APPLICATION NUMBER: US/10/505,263
; PRIOR FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: US 60/360,275
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US03/06469
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 96
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-263-96

Query Match 33.8%; Score 753.5; DB 6; Length 739;
Best Local Similarity 35.6%; Pred. No. 7.2e-52;
Matches 149; Conservative 98; Mismatches 159; Indels 13; Gaps 2;

Qy 32 LPILDWAPHYNLKENLLPDTVSGIMLAVQVQVTLQGLAFVLSVHPVGLYSLPAIIYA 91
Db 93 LPVLQMLPKYDLKKNILGDMVMSGLIVGILLVPOSIAYSLLAGOEPYGLYTSFFASIIYF 152
Qy 92 IFGMGHVATGTALTSLISANAVRIVPQ-----NMQNLTQSNSTVL-----GLS 138
Db 153 LLGTSRHSIVGIFGVLCIMIGETVDRBLQKAGYDNAHSAPSLGMVNSGSTLNLNHTSDRIC 212
Qy 139 DFEMQRHVAANAASVFLGGVIVQAMFVLQGSATFVVTEPVIAMTGAATHVVTQVKYL 198
Db 213 DKSCYALMVGSTVTFIAGVYQVAMGPFQGVSVVLSDDLGSFTLTSGAKYLL 272
Qy 199 LGMKMPYISGLPFPFYIYAVFENIKSVRLLEALLSLLSVLVVLKELNEQFKRKIKV 258
Db 273 LGLNLPRTNGVSLITTWIHFVRNIHKNLCLDLITSLCLLVLPSEKELNEHFKSLKAP 332
Qy 259 LPVDLVLIIAASPACYCTNMNTYGLVGVGHIPQIPSPRAPMNLISAVITEAGVALV 318
Db 333 IPLELVVVVAATLASHFGKLGNGYNSIAGHPTGFMPPKAPDWSLIPNVAIDAISII 392
Qy 319 GYVASLALAAQSKAKPKYISIDNQBFLAHLGNISVSSPPFCIPSAAMGRTAGLYSTGAK 378
Db 393 GPATVVSLSSEMFACKGCHGYTVKANQEMYAIQFCNIIIPSPFHCTTSAALAKTLVKESTGCH 452
Qy 379 TQVACLISCIFVLIYIYVAIGPLLYWLMPCVLASIIIVVGLKGMILQFRDLKKYWNVDKID 437
Db 453 TQLSGVVVVALVLLVLLVIAPLFYSLQKCVLGVITIVNLRGALLKFRDLPKMWRLSRMD 511

RESULT 4
US-10-505-263-87
; Sequence 87, Application US/10505263
; Publication No. US20060014940A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Case Western Reserve University
; APPLICANT: The Brigham and Women's Hospital, Inc.
; APPLICANT: Mount, David B
; APPLICANT: Romero, Michael

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2006, 18:48:48 ; Search time 168 Seconds
(without alignments)
1091.828 Million cell updates/sec

Title: US-10-761-033-2_COPY_1_439

Perfect score: 2232

Sequence: 1 MTGAKRKKKSLMSKMTTPQ.....MLIQRLDKKYNVDKIDWG 439

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA Main:
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------|
| 1 | 2232 | 100.0 | 598 | 3 | US-09-875-811-12 |
| 2 | 2232 | 100.0 | 598 | 4 | US-10-761-033-12 |
| 3 | 2232 | 100.0 | 605 | 3 | US-09-875-811-8 |
| 4 | 2232 | 100.0 | 605 | 4 | US-10-761-033-8 |
| 5 | 2232 | 100.0 | 621 | 3 | US-09-875-811-4 |
| 6 | 2232 | 100.0 | 621 | 4 | US-10-761-033-4 |
| 7 | 2232 | 100.0 | 656 | 3 | US-09-875-811-10 |
| 8 | 2232 | 100.0 | 656 | 4 | US-10-761-033-10 |
| 9 | 2232 | 100.0 | 663 | 3 | US-09-875-811-6 |
| 10 | 2232 | 100.0 | 663 | 4 | US-10-761-033-6 |
| 11 | 2232 | 100.0 | 663 | 3 | US-10-222-009-14 |
| 12 | 2232 | 100.0 | 663 | 4 | US-10-761-033-6 |
| 13 | 2232 | 100.0 | 679 | 3 | US-09-875-811-2 |
| 14 | 2232 | 100.0 | 679 | 4 | US-10-761-033-2 |
| 15 | 2227 | 99.8 | 656 | 4 | US-10-222-009-13 |
| 16 | 2077 | 93.1 | 656 | 5 | US-10-505-316-2 |
| 17 | 2077 | 93.1 | 656 | 5 | US-10-505-316-4 |
| 18 | 2064.5 | 92.5 | 671 | 4 | US-10-415-378-4 |
| 19 | 938.5 | 42.0 | 431 | 5 | US-10-450-763-36658 |
| 20 | 825.5 | 37.0 | 555 | 4 | US-10-478-758-6 |
| 21 | 753.5 | 33.8 | 739 | 4 | US-10-205-823-369 |
| 22 | 753.5 | 33.8 | 739 | 4 | US-10-328-194A-3 |
| 23 | 753.5 | 33.8 | 739 | 4 | US-10-736-461-1 |
| 24 | 753.5 | 33.8 | 739 | 6 | US-11-051-454-369 |
| 25 | 747.5 | 33.5 | 780 | 5 | US-10-631-467-3 |
| 26 | 747.5 | 33.5 | 780 | 5 | US-10-631-467-667 |
| 27 | 747.5 | 33.5 | 790 | 4 | US-10-295-027-930 |

| | | | | | | |
|----|-------|------|------|---|--------------------|-------------------|
| 28 | 741.5 | 33.2 | 764 | 3 | US-09-981-353-73 | Sequence 73, Appl |
| 29 | 741.5 | 33.2 | 764 | 4 | US-10-235-994-20 | Sequence 20, Appl |
| 30 | 732 | 32.8 | 714 | 3 | US-09-749-589-4 | Sequence 4, Appl |
| 31 | 732 | 32.8 | 714 | 5 | US-10-684-532-4 | Sequence 4, Appl |
| 32 | 721.5 | 32.3 | 751 | 3 | US-09-795-633-14 | Sequence 14, Appl |
| 33 | 721.5 | 32.3 | 751 | 4 | US-10-156-239-14 | Sequence 14, Appl |
| 34 | 721.5 | 32.3 | 751 | 4 | US-10-199-485-14 | Sequence 14, Appl |
| 35 | 721.5 | 32.3 | 780 | 4 | US-10-420-495-11 | Sequence 11, Appl |
| 36 | 721.5 | 32.3 | 4115 | 4 | US-10-038-854-4 | Sequence 4, Appl |
| 37 | 718.5 | 32.2 | 780 | 5 | US-10-631-467-14 | Sequence 14, Appl |
| 38 | 718.5 | 32.2 | 780 | 5 | US-10-631-467-1456 | Sequence 1456, Ap |
| 39 | 708 | 31.7 | 744 | 4 | US-10-420-495-1 | Sequence 1, Appl |
| 40 | 708 | 31.7 | 744 | 4 | US-10-420-495-3 | Sequence 3, Appl |
| 41 | 702 | 31.5 | 505 | 4 | US-10-467-685-4 | Sequence 4, Appl |
| 42 | 662.5 | 29.7 | 790 | 5 | US-10-505-316-10 | Sequence 10, Appl |
| 43 | 645.5 | 28.9 | 753 | 3 | US-09-749-589-2 | Sequence 2, Appl |
| 44 | 645.5 | 28.9 | 753 | 5 | US-10-684-532-2 | Sequence 2, Appl |
| 45 | 645.5 | 28.9 | 791 | 5 | US-10-505-316-6 | Sequence 6, Appl |

ALIGNMENTS

RESULT 1

US-09-875-811-12
; Sequence 12, Application US/09875811
; Patent No. US20020032321A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; TITLE OF INVENTION: No. US20020032321A1el Human Transporter Proteins and Polynucleotic
; FILE OF INVENTION: Same
; FILE REFERENCE: LEX-0186-USA
; CURRENT APPLICATION NUMBER: US/09/875,811
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/210,045
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 598
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-875-811-12

Query Match

100.0%; Score 2232; DB 3; Length 598;

Best Local Similarity 100.0%; Pred. No. 8.5e-201; Mismatches 0; Indels 0; Gaps 0;
Matches 439; Conservative 0

| | | | | |
|----|-----|------------------------|---|-----|
| Qy | 1 | MTGAKRKKKSLMSKMTTPQCE | IIQMCRRRLPILDWAPHYLNKENLLPDTVSGIMLAVQ | 60 |
| Db | 1 | MTGAKRKKKSLMSKMTTPQCE | IIQMCRRRLPILDWAPHYLNKENLLPDTVSGIMLAVQ | 60 |
| Qy | 61 | QVTGGLAFALVSSVHPFGLYGS | LPFPAIIYAFGNGHHVATGTALTSLISANAVRIVP | 120 |
| Db | 61 | QVTGGLAFALVSSVHPFGLYGS | LPFPAIIYAFGNGHHVATGTALTSLISANAVRIVP | 120 |
| Qy | 121 | QNMQLTQSTNSVLGLSDFERQ | RIHVAAAASFVGIVQAMFVLQGSATFVWTEPVIS | 180 |
| Db | 121 | QNMQLTQSTNSVLGLSDFERQ | RIHVAAAASFVGIVQAMFVLQGSATFVWTEPVIS | 180 |
| Qy | 181 | AMTTGAATHVVTQVKYLLGKM | PFYISGPLGFFYYIAYVVFENIKSVRLLEALLLSLSIW | 240 |
| Db | 181 | AMTTGAATHVVTQVKYLLGKM | PFYISGPLGFFYYIAYVVFENIKSVRLLEALLLSLSIW | 240 |
| Qy | 241 | LVLVKELNEQPKRKIKVVL | PVDLVLIITAAFPACVTNNNTYGLGVGHIPOGISPRAP | 300 |
| Db | 241 | LVLVKELNEQPKRKIKVVL | PVDLVLIITAAFPACVTNNNTYGLGVGHIPOGISPRAP | 300 |
| Qy | 301 | PMNLSAVITEAFGVALGVY | VASLALAQGSAAKKPKYSIDDDNQEFLAHGLSNIVSSPFPCCI | 360 |
| Db | 301 | PMNLSAVITEAFGVALGVY | VASLALAQGSAAKKPKYSIDDDNQEFLAHGLSNIVSSPFPCCI | 360 |

Qy 361 PSAAAMGRTAGLYSTGAKTQVACLISCIFVLIVIIYAIGPLLYWLPVCVLASIIIVVGLKGM 420
Db 361 PSAAAMGRTAGLYSTGAKTQVACLISCIFVLIVIIYAIGPLLYWLPVCVLASIIIVVGLKGM 420
Qy 421 LIQPRDLKKYNNVDKIDWG 439
Db 421 LIQPRDLKKYNNVDKIDWG 439

RESULT 2

US-10-761-033-12
; Sequence 12, Application US/10761033
; Publication No. US20040138442A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; TITLE OF INVENTION: Novel Human Transporter Proteins and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0186-USA
; CURRENT APPLICATION NUMBER: US/10/761,033
; CURRENT FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US/09/875,811
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/210,045
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 598
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-761-033-12

Query Match 100.0%; Score 2232; DB 4; Length 598;
Best Local Similarity 100.0%; Pred. No. 8.5e-201;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTGAKRKKKSLMWSKMTPOCEDIIQWCRRLPILDWAPHYLNKENTLPDVTSGIMLAVQ 60
Db 1 MTGAKRKKKSLMWSKMTPOCEDIIQWCRRLPILDWAPHYLNKENTLPDVTSGIMLAVQ 60
Qy 61 QVTOGLAPAVLSSVHPVFLYGLSLPFAIIYIAIFGMGHHVATGTALTSLISANAVRIVP 120
Db 61 QVTOGLAPAVLSSVHPVFLYGLSLPFAIIYIAIFGMGHHVATGTALTSLISANAVRIVP 120
Qy 121 QNMQLTTQSNSTSVLGLSDPEMORIHVAAVSLGGVIVQAMFVLQGSATFVVTPEVIS 180
Db 121 QNMQLTTQSNSTSVLGLSDPEMORIHVAAVSLGGVIVQAMFVLQGSATFVVTPEVIS 180
Qy 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYIYAVFENIKSVRLLEALLLSLSIV 240
Db 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYIYAVFENIKSVRLLEALLLSLSIV 240
Qy 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASPCYCTNMNTYGLVGVGHIPOGIPSPRAP 300
Db 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASPCYCTNMNTYGLVGVGHIPOGIPSPRAP 300
Qy 301 PMNLSAVITEAFGVALVGYVASLALAQSAKPKYSIDDDQEFPLAHGLSNIVSFPFCI 360
Db 301 PMNLSAVITEAFGVALVGYVASLALAQSAKPKYSIDDDQEFPLAHGLSNIVSFPFCI 360
Qy 361 PSAAAMGRTAGLYSTGAKTQVACLISCIFVLIVIIYAIGPLLYWLPVCVLASIIIVVGLKGM 420
Db 361 PSAAAMGRTAGLYSTGAKTQVACLISCIFVLIVIIYAIGPLLYWLPVCVLASIIIVVGLKGM 420
Qy 421 LIQPRDLKKYNNVDKIDWG 439
Db 421 LIQPRDLKKYNNVDKIDWG 439

RESULT 3

US-09-875-811-8
; Sequence 8, Application US/09875811

; Patent No. US20020032321A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; TITLE OF INVENTION: No. US20020032321A1el Human Transporter Proteins and Polynucleotides
; FILE REFERENCE: LEX-0186-USA
; CURRENT APPLICATION NUMBER: US/09/875,811
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/210,045
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 605
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-875-811-8

Query Match 100.0%; Score 2232; DB 3; Length 605;
Best Local Similarity 100.0%; Pred. No. 8.7e-201;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTGAKRKKKSLMWSKMTPOCEDIIQWCRRLPILDWAPHYLNKENTLPDVTSGIMLAVQ 60
Db 1 MTGAKRKKKSLMWSKMTPOCEDIIQWCRRLPILDWAPHYLNKENTLPDVTSGIMLAVQ 60
Qy 61 QVTOGLAPAVLSSVHPVFLYGLSLPFAIIYIAIFGMGHHVATGTALTSLISANAVRIVP 120
Db 61 QVTOGLAPAVLSSVHPVFLYGLSLPFAIIYIAIFGMGHHVATGTALTSLISANAVRIVP 120
Qy 121 QNMQLTTQSNSTSVLGLSDPEMORIHVAAVSLGGVIVQAMFVLQGSATFVVTPEVIS 180
Db 121 QNMQLTTQSNSTSVLGLSDPEMORIHVAAVSLGGVIVQAMFVLQGSATFVVTPEVIS 180
Qy 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYIYAVFENIKSVRLLEALLLSLSIV 240
Db 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYIYAVFENIKSVRLLEALLLSLSIV 240
Qy 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASPCYCTNMNTYGLVGVGHIPOGIPSPRAP 300
Db 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASPCYCTNMNTYGLVGVGHIPOGIPSPRAP 300
Qy 301 PMNLSAVITEAFGVALVGYVASLALAQSAKPKYSIDDDQEFPLAHGLSNIVSFPFCI 360
Db 301 PMNLSAVITEAFGVALVGYVASLALAQSAKPKYSIDDDQEFPLAHGLSNIVSFPFCI 360
Qy 361 PSAAAMGRTAGLYSTGAKTQVACLISCIFVLIVIIYAIGPLLYWLPVCVLASIIIVVGLKGM 420
Db 361 PSAAAMGRTAGLYSTGAKTQVACLISCIFVLIVIIYAIGPLLYWLPVCVLASIIIVVGLKGM 420
Qy 421 LIQPRDLKKYNNVDKIDWG 439
Db 421 LIQPRDLKKYNNVDKIDWG 439

RESULT 4

US-10-761-033-8
; Sequence 8, Application US/10761033
; Publication No. US20040138442A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; TITLE OF INVENTION: Novel Human Transporter Proteins and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0186-USA
; CURRENT APPLICATION NUMBER: US/10/761,033
; CURRENT FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US/09/875,811
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/210,045
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 13

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2006, 18:48:03 ; Search time 46 Seconds
(without alignments)
789.014 Million cell updates/sec

Title: US-10-761-033-2_COPY_1_439

Perfect score: 2232
Sequence: 1 MTGAKRKKKSMKMTPOQEDIIQWCRRLPILDWAPHYLNKKNLLPDTVSGIMLA VQ 439

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*

- 1: /cgn2_6/ptodata/1/1aa/5_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/6_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/H_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/PCUS_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/RE_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 2232 | 100.0 | 598 | 2 | US-09-875-811-12 |
| 2 | 2232 | 100.0 | 605 | 2 | US-09-875-811-8 |
| 3 | 2232 | 100.0 | 621 | 2 | US-09-875-811-4 |
| 4 | 2232 | 100.0 | 656 | 2 | US-09-875-811-10 |
| 5 | 2232 | 100.0 | 663 | 2 | US-09-875-811-6 |
| 6 | 2232 | 100.0 | 679 | 2 | US-09-875-811-2 |
| 7 | 753.5 | 33.8 | 748 | 2 | US-09-949-016-10387 |
| 8 | 747.5 | 33.5 | 803 | 2 | US-09-949-016-11498 |
| 9 | 741.5 | 33.2 | 764 | 1 | US-08-424-567-2 |
| 10 | 741.5 | 33.2 | 764 | 1 | US-08-711-928-2 |
| 11 | 741.5 | 33.2 | 764 | 2 | US-09-184-937-2 |
| 12 | 741.5 | 33.2 | 790 | 2 | US-09-949-016-11220 |
| 13 | 721.5 | 32.3 | 780 | 2 | US-09-785-381-11 |
| 14 | 708 | 31.7 | 744 | 2 | US-09-785-381-1 |
| 15 | 708 | 31.7 | 744 | 2 | US-09-785-381-3 |
| 16 | 472 | 21.1 | 593 | 2 | US-09-720-317A-22 |
| 17 | 455.5 | 20.4 | 621 | 2 | US-09-720-317A-16 |
| 18 | 453 | 20.3 | 970 | 2 | US-09-795-927-7 |
| 19 | 451.5 | 20.2 | 646 | 2 | US-09-720-317A-28 |
| 20 | 444.5 | 19.9 | 680 | 2 | US-09-720-317A-18 |
| 21 | 443.5 | 19.9 | 658 | 2 | US-09-720-317A-24 |
| 22 | 423 | 19.0 | 685 | 2 | US-09-720-317A-31 |
| 23 | 421 | 18.9 | 295 | 2 | US-09-785-381-5 |
| 24 | 418.5 | 18.8 | 646 | 2 | US-09-720-317A-26 |
| 25 | 414 | 18.5 | 579 | 2 | US-09-720-317A-4 |
| 26 | 412.5 | 18.5 | 660 | 2 | US-09-720-317A-23 |
| 27 | 408 | 18.3 | 688 | 2 | US-09-720-317A-2 |

| | | | | | |
|----|-------|------|-----|---|----------------------|
| 28 | 395 | 17.7 | 656 | 2 | US-09-720-317A-20 |
| 29 | 389.5 | 17.5 | 590 | 2 | US-09-902-540-14944 |
| 30 | 389.5 | 17.5 | 631 | 2 | US-09-720-317A-29 |
| 31 | 388 | 17.4 | 644 | 2 | US-09-720-317A-25 |
| 32 | 383 | 17.2 | 596 | 2 | US-09-252-991A-23812 |
| 33 | 376 | 16.8 | 660 | 2 | US-09-720-317A-30 |
| 34 | 351 | 15.7 | 585 | 2 | US-09-602-787A-616 |
| 35 | 322 | 14.4 | 828 | 2 | US-09-248-796A-20746 |
| 36 | 306 | 13.7 | 616 | 2 | US-09-543-681A-4421 |
| 37 | 297 | 13.3 | 566 | 2 | US-09-543-681A-4544 |
| 38 | 282.5 | 12.7 | 533 | 2 | US-09-107-532A-5244 |
| 39 | 282.5 | 12.7 | 537 | 2 | US-09-602-787A-618 |
| 40 | 271 | 12.1 | 562 | 2 | US-09-489-039A-10405 |
| 41 | 263 | 11.8 | 436 | 2 | US-09-134-000C-3950 |
| 42 | 263 | 11.8 | 535 | 2 | US-09-252-991A-21805 |
| 43 | 255.5 | 11.4 | 483 | 2 | US-09-710-279-3132 |
| 44 | 250 | 11.2 | 575 | 2 | US-09-438-185A-1015 |
| 45 | 243 | 10.9 | 151 | 2 | US-09-270-767-33024 |

ALIGNMENTS

RESULT 1

US-09-875-811-12
; Sequence 12, Application US/09875811
; Patent No. 6703495
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; TITLE OF INVENTION: No. 6703495el Human Transporter Proteins and Polynucleotides Encod
; FILE REFERENCE: LEX-0186-USA
; CURRENT APPLICATION NUMBER: US/09/875,811
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/210,045
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 598
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-875-811-12

Query Match 100.0%; Score 2232; DB 2; Length 598;
Best Local Similarity 100.0%; Pred. No. 2.1e-215;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| Qy | 1 | MTGAKRKKKSMKMTPOQEDIIQWCRRLPILDWAPHYLNKKNLLPDTVSGIMLA VQ | 60 |
| Db | 1 | MTGAKRKKKSMKMTPOQEDIIQWCRRLPILDWAPHYLNKKNLLPDTVSGIMLA VQ | 60 |
| Qy | 61 | QVTOGLAPAVLSSVHPVFGYLSLFPALIIAIFGKHVATGTALTSLISANAVR LVP | 120 |
| Db | 61 | QVTOGLAPAVLSSVHPVFGYLSLFPALIIAIFGKHVATGTALTSLISANAVR LVP | 120 |
| Qy | 121 | QNMQLTQNTSVLGLSDPEMQRIHVAAAVSFLGGVITQVAMFVLQGSATFVVT E PVIS | 180 |
| Db | 121 | QNMQLTQNTSVLGLSDPEMQRIHVAAAVSFLGGVITQVAMFVLQGSATFVVT E PVIS | 180 |
| Qy | 181 | AMTTGAATHVVTQVKKLLGKMPYISGPIGFFYIYAVFENIKSVRLRALLLSL SVV | 240 |
| Db | 181 | AMTTGAATHVVTQVKKLLGKMPYISGPIGFFYIYAVFENIKSVRLRALLLSL SVV | 240 |
| Qy | 241 | LVLVKELNEQPKRKIKVLPVDLVLIIAASPACTNMENTYGLVGVCHIQGGIPSP RAP | 300 |
| Db | 241 | LVLVKELNEQPKRKIKVLPVDLVLIIAASPACTNMENTYGLVGVCHIQGGIPSP RAP | 300 |
| Qy | 301 | PMNLTSAVITAEFGVALGVYVASLALAQGSACKPKYSIDDDNQEFHLAHLGSLNIVSFFFCI | 360 |
| Db | 301 | PMNLTSAVITAEFGVALGVYVASLALAQGSACKPKYSIDDDNQEFHLAHLGSLNIVSFFFCI | 360 |

QY 361 PSAAAMGRTAGLYTGAKTQVACLISCIFVLIVIIYAIGPLLYLWLPVCVLASIIIVVGLKGM 420
Db ||||||||||||||||||
QY 421 LIQFRDLKKYWNVDKIDWG 439
Db ||||||||||||||||||
QY 421 LIQFRDLKKYWNVDKIDWG 439

RESULT 2

US-09-875-811-8
; Sequence 8, Application US/09875811
; Patent No. 6703495
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; TITLE OF INVENTION: No. 6703495el Human Transporter Proteins and Polynucleotides Encod
; FILE REFERENCE: LEX-0186-USA
; CURRENT APPLICATION NUMBER: US/09/875,811
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/210,045
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 605
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-875-811-8

Query Match 100.0%; Score 2232; DB 2; Length 605;
Best Local Similarity 100.0%; Pred. No. 2.1e-215;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTGAKRKKKSLMSKQHTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAQV 60
Db 1 MTGAKRKKKSLMSKQHTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAQV 60

QY 61 QVTQGLAPAVLSSVHPVGLYGLSLFPALIIYAIFGHHVATGTALTSLISANAVERIVP 120
Db 61 QVTQGLAPAVLSSVHPVGLYGLSLFPALIIYAIFGHHVATGTALTSLISANAVERIVP 120
QY 121 QNMQLTTQSNLTSVLGLSDPFEMQRIHVAAAVSFLGGVIVQAMFVLQGSATFVVTPEVIS 180
Db 121 QNMQLTTQSNLTSVLGLSDPFEMQRIHVAAAVSFLGGVIVQAMFVLQGSATFVVTPEVIS 180
QY 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYIYAVPENIKSVRLKALLLSLSIV 240
Db 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYIYAVPENIKSVRLKALLLSLSIV 240
QY 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVVGHIPOGIPSPRAP 300
Db 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVVGHIPOGIPSPRAP 300
QY 301 PMNILSAVITEAFGVALGVYASLALAQSAKKPKYSIDDNQBEFLAHLGSLNIVSSPPFCI 360
Db 301 PMNILSAVITEAFGVALGVYASLALAQSAKKPKYSIDDNQBEFLAHLGSLNIVSSPPFCI 360
QY 361 PSAAAMGRTAGLYTGAKTQVACLISCIFVLIVIIYAIGPLLYLWLPVCVLASIIIVVGLKGM 420
Db 361 PSAAAMGRTAGLYTGAKTQVACLISCIFVLIVIIYAIGPLLYLWLPVCVLASIIIVVGLKGM 420
QY 421 LIQFRDLKKYWNVDKIDWG 439
Db 421 LIQFRDLKKYWNVDKIDWG 439

RESULT 3

US-09-875-811-4
; Sequence 4, Application US/09875811
; Patent No. 6703495
; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; TITLE OF INVENTION: No. 6703495el Human Transporter Proteins and Polynucleotides Encod
; FILE REFERENCE: LEX-0186-USA
; CURRENT APPLICATION NUMBER: US/09/875,811
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/210,045
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 621
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-875-811-4

Query Match 100.0%; Score 2232; DB 2; Length 621;
Best Local Similarity 100.0%; Pred. No. 2.2e-215;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTGAKRKKKSLMSKQHTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAQV 60
Db 1 MTGAKRKKKSLMSKQHTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAQV 60
QY 61 QVTQGLAPAVLSSVHPVGLYGLSLFPALIIYAIFGHHVATGTALTSLISANAVERIVP 120
Db 61 QVTQGLAPAVLSSVHPVGLYGLSLFPALIIYAIFGHHVATGTALTSLISANAVERIVP 120
QY 121 QNMQLTTQSNLTSVLGLSDPFEMQRIHVAAAVSFLGGVIVQAMFVLQGSATFVVTPEVIS 180
Db 121 QNMQLTTQSNLTSVLGLSDPFEMQRIHVAAAVSFLGGVIVQAMFVLQGSATFVVTPEVIS 180
QY 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYIYAVPENIKSVRLKALLLSLSIV 240
Db 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYIYAVPENIKSVRLKALLLSLSIV 240
QY 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVVGHIPOGIPSPRAP 300
Db 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVVGHIPOGIPSPRAP 300
QY 301 PMNILSAVITEAFGVALGVYASLALAQSAKKPKYSIDDNQBEFLAHLGSLNIVSSPPFCI 360
Db 301 PMNILSAVITEAFGVALGVYASLALAQSAKKPKYSIDDNQBEFLAHLGSLNIVSSPPFCI 360
QY 361 PSAAAMGRTAGLYTGAKTQVACLISCIFVLIVIIYAIGPLLYLWLPVCVLASIIIVVGLKGM 420
Db 361 PSAAAMGRTAGLYTGAKTQVACLISCIFVLIVIIYAIGPLLYLWLPVCVLASIIIVVGLKGM 420
QY 421 LIQFRDLKKYWNVDKIDWG 439
Db 421 LIQFRDLKKYWNVDKIDWG 439

RESULT 4

US-09-875-811-10
; Sequence 10, Application US/09875811
; Patent No. 6703495
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; TITLE OF INVENTION: No. 6703495el Human Transporter Proteins and Polynucleotides Encod
; FILE REFERENCE: LEX-0186-USA
; CURRENT APPLICATION NUMBER: US/09/875,811
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/210,045
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 656
; TYPE: PRT

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2006, 18:44:08 ; Search time 40 seconds
(without alignments)
1055.979 Million cell updates/sec

Title: US-10-761-033-2_COPY_1_439

Perfect score: 2232
Sequence: 1 MTGAKRKKXKMLMSRMTQP.....MLIQPRDLKKYWNVDKIDWG 439

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_80.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 752.5 | 33.7 | 739 | 2 A54808 | diastrophic dyspla |
| 2 | 741.5 | 33.2 | 764 | 2 A47456 | down-regulated in |
| 3 | 668.5 | 30.0 | 703 | 2 A49994 | sulfate transporte |
| 4 | 648 | 29.0 | 611 | 2 T27820 | hypothetical prote |
| 5 | 638.5 | 28.6 | 700 | 2 T23629 | hypothetical prote |
| 6 | 611 | 27.4 | 749 | 2 T23628 | hypothetical prote |
| 7 | 571.5 | 25.6 | 758 | 2 T25751 | hypothetical prote |
| 8 | 553.5 | 24.8 | 380 | 2 T26165 | hypothetical prote |
| 9 | 543 | 24.3 | 650 | 2 T16077 | hypothetical prote |
| 10 | 496.5 | 22.2 | 750 | 2 B90137 | sulfate permease |
| 11 | 486.5 | 21.8 | 809 | 2 T32945 | hypothetical prote |
| 12 | 451.5 | 20.2 | 646 | 2 T01079 | sulfate transport |
| 13 | 447.5 | 20.0 | 605 | 2 T70479 | high affinity sulf |
| 14 | 443.5 | 19.9 | 658 | 2 T49069 | sulfate transporte |
| 15 | 443.5 | 19.9 | 658 | 2 T48901 | sulfate transporte |
| 16 | 443.5 | 19.9 | 703 | 2 T51161 | hypothetical prote |
| 17 | 418.5 | 18.8 | 646 | 2 T48902 | sulfate transporte |
| 18 | 412.5 | 18.5 | 553 | 2 A11498 | probable sulfate t |
| 19 | 409.5 | 18.3 | 553 | 2 AB1140 | probable sulfate t |
| 20 | 409.5 | 18.3 | 662 | 2 S51764 | sulfate transporte |
| 21 | 403.5 | 18.1 | 667 | 2 S51763 | sulfate transporte |
| 22 | 393.5 | 17.6 | 631 | 2 B86365 | probable sulphate |
| 23 | 388 | 17.4 | 644 | 2 S51765 | sulfate transport |
| 24 | 386 | 17.3 | 660 | 2 T04416 | sulfate transport |
| 25 | 372 | 16.7 | 566 | 2 S74633 | high affinity sulf |
| 26 | 371.5 | 16.6 | 573 | 2 D83440 | probable sulfate t |
| 27 | 369.5 | 16.6 | 677 | 2 T50022 | sulfate transporte |
| 28 | 357 | 16.0 | 649 | 2 T51839 | sulfate transport |
| 29 | 357 | 16.0 | 649 | 2 T00946 | probable sulfate t |

ALIGNMENTS

RESULT 1

A54808

diastrophic dysplasia-associated sulfate transport protein - human

C:Species: Homo sapiens (man)

C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 07-May-1999

C:Accession: A54808

R:Haestbacka, J.; de la Chapelle, A.; Mahtani, M.M.; Clines, G.; Reeve-Daly, M.P.; Daly,

der, E.S. 1073-1087, 1994

Cell, 78, 1073-1087, 1994

A:Title: The diastrophic dysplasia gene encodes a novel sulfate transporter: positional c

A:Reference number: A54808; MUID:95007757; PMID:7923357

A:Accession: A54808

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-739 <HAB>

A:Cross-references: UNIPARC:UPI0000178906; GB:U14528

C:Genetics:

A:Gene: GDB:DTD; DTDST

A:Cross-references: GDB:125421; OMIM:222600

A:Map position: Sq32-Sq33.1

C:Superfamily: sulfate transport protein

Query Match 33.7%; Score 752.5; DB 2; Length 739;

Best Local Similarity 35.6%; Pred. No. 8.9e-51;

Matches 149; Conservative 98; Mismatches 159; Indels 13; Gaps 2;

| | | | |
|----|-----|--|-----|
| Qy | 32 | LPILDWAPHNLKENLLPDTVSGIMLAVQVQTQGLAFVLSVHPVGLYGLSPALIYA | 91 |
| Db | 93 | LPVLQMLPKYDLKKNILGDMVSLIVGILVPOSIAVSLLAGQBPVYGLYTSFPASIIYF | 152 |
| Qy | 92 | IFGMGHVATGTPTALTSLSANAVRIVPO-----NMQNLTTQSNLSVL-----GLS | 138 |
| Db | 153 | LLGTSRAISVIGIFVLCIMGETVDRELQAGYDHAHSAPSLGMVNSGSLNHTSDRIC | 212 |
| Qy | 139 | DFEMQRIHVAANVSLFGGVIVQAMFVLQLSATPVVTPVVISAMTTGAATHVTSQVKYL | 198 |
| Db | 213 | DKSCYALMVGSTVTFIAGVTVQAMGFQGVSVLSALLSGFVTGASPTILTSQAKYL | 272 |
| Qy | 199 | IGMKNPISGLPFPFYIYAVFENIKSVRLKALLLSISIVLVVKELNEQPKRKIKV | 258 |
| Db | 273 | LGLMLPRTNGVSLTITTHVFRNIHKTNIICDLTSLCLLVLLPTKELNHFHKSCLKAP | 332 |
| Qy | 259 | LPVDVLIIAASPACYCTNMNTYGLFVWGHIPGIPSPRAPPNNILSAVITFAFGVALV | 318 |
| Db | 333 | IPIELVVVVVAATLASHFGKLUHNTYNSIAGHIPGFMPPKVPENWLIPSAVDAIAISII | 392 |
| Qy | 319 | GYVASLALAQSAKKPKYISIDDDNQEFPLAHGLSNITVSSFFFCIPSAAMGRTAGLYSTGAK | 378 |
| Db | 393 | GFAITVLSLSEMFACKHGYTVKANQEVVAIGFCNIIIPSFHCFHTTSAALAKTLVKESTGCH | 452 |
| Qy | 379 | TQVACLISCFVLIVIVAIGPGLLWLPMPCVLASIIIVGLKGLIQPRDLKKYWNVDKID | 437 |

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2006, 18:40:53 ; Search time 232 Seconds
(without alignments)
1335.030 Million cell updates/sec

Title: US-10-761-033-2_COPY_1_439

Perfect score: 2232

Sequence: 1 MTGAKRKKKSLMSKMTTPQ.....MLIQFLDLKKYNNVDKIDWG 439

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 2232 | 100.0 | 656 | 2 | Q96RN2 HUMAN |
| 2 | 2232 | 100.0 | 663 | 2 | Q8TE53 homo sapien |
| 3 | 2227 | 99.8 | 656 | 2 | Q8TE54 HUMAN |
| 4 | 2208 | 98.9 | 656 | 2 | Q8RAL2 PONY |
| 5 | 2077 | 93.1 | 656 | 2 | Q8R223 MOUSE |
| 6 | 812.5 | 36.4 | 711 | 2 | Q5GLK3 CIOIN |
| 7 | 811.5 | 36.4 | 492 | 2 | Q5EB10 MOUSE |
| 8 | 805 | 36.1 | 735 | 2 | Q5YLVS XENLA |
| 9 | 800 | 35.8 | 585 | 2 | Q6GQ18 XENLA |
| 10 | 800 | 35.8 | 720 | 2 | Q4V7Y0 XENLA |
| 11 | 800 | 35.8 | 754 | 2 | Q8AYP0 ANGJA |
| 12 | 797.5 | 35.7 | 713 | 2 | Q6L968 ANGJA |
| 13 | 789 | 35.3 | 766 | 2 | Q6XDT1 CIOIN |
| 14 | 785 | 35.2 | 739 | 2 | Q7T2C4 BRARE |
| 15 | 781 | 35.0 | 726 | 2 | Q6DJN0 XENLA |
| 16 | 774 | 34.7 | 739 | 1 | S26A2 MOUSE |
| 17 | 774 | 34.7 | 739 | 2 | Q8R213 MOUSE |
| 18 | 774 | 34.7 | 739 | 2 | Q543D6 mouse |
| 19 | 772 | 34.6 | 714 | 2 | Q7T2N6 BRARE |
| 20 | 765.5 | 34.3 | 804 | 2 | Q6L967 ANGJA |
| 21 | 763.5 | 34.2 | 563 | 2 | Q8NGQ4 HUMAN |
| 22 | 762.5 | 34.2 | 757 | 2 | Q8BGH1 RABIT |
| 23 | 759 | 34.0 | 629 | 1 | S26A2 RAT |
| 24 | 757 | 33.9 | 729 | 2 | Q4SVG1 TETNG |
| 25 | 755.5 | 33.8 | 734 | 2 | Q8BEG8 BOVIN |
| 26 | 753.5 | 33.8 | 739 | 1 | S26A2 HUMAN |
| 27 | 750.5 | 33.6 | 736 | 2 | Q65AC2 HORSE |
| 28 | 748.5 | 33.5 | 595 | 2 | Q8RGV3 BRARE |
| 29 | 747.5 | 33.5 | 780 | 1 | PEND HUMAN |
| 30 | 743.5 | 33.3 | 733 | 2 | Q69DJ1 BUBBU |
| 31 | 743.5 | 33.3 | 780 | 2 | Q5NVA6 PONGY |

32 742.5 33.3 713 2 Q5U500 XENLA
33 742.5 33.3 719 2 Q6XQ07 XENLA
34 741.5 33.2 764 1 S26A3 HUMAN
35 739 33.1 651 2 Q6MZW7 HUMAN
36 738 33.1 757 1 S26A3 MOUSE
37 735.5 33.0 734 2 Q8GYJ3 SHEEP
38 735 32.9 727 2 Q5R7H3 PONY
39 734 32.9 753 2 Q5GM09 FIG
40 732 32.8 738 2 Q9NQ11 HUMAN
41 732 32.8 758 2 Q548A7 HUMAN
42 732 32.8 759 1 S26A6 HUMAN
43 726.5 32.5 735 2 Q812E2 MOUSE
44 726 32.5 757 1 S26A3 RAT
45 725.5 32.5 735 2 Q8CJD0 MOUSE

ALIGNMENTS

RESULT 1
Q96RN2 HUMAN
ID Q56RN2 HUMAN PRELIMINARY; PRT; 656 AA.
AC Q96RN2
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DB Anion transporter (SLC26A7 protein).
GN Name=SLC26A7; Synonyms=SUT2/SLC26A7;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21950770; PubMed=11834742; DOI=10.1074/jbc.M11802200;
RA Lohi H., Kujala M., Makela S., Lehtonen E., Kestila M.,
Saarialho-Kere U., Markovich D., Kere J.;
RT "Functional characterization of three novel tissue-specific anion
exchangers SLC26A7, -A8, and -A9".
RL J. Biol. Chem. 277.14246-14254(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=21818883; PubMed=11829495; DOI=10.1006/geno.2002.6689;
RA Vincourt J.B., Jullien D., Kossida S., Amalric P., Girard J.P.;
RT "Molecular cloning of SLC26A7, a novel member of the SLC26
sulfate/anion transporter family, from high endothelial venules and
kidney".
RL Genomics 79:249-256(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Girard J.;
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins L., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Frange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Kidney;
RC TISSUE=Placenta;
RG NIH MGC Project;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF331521; AAK9565.1; -; mRNA.
DR EMBL; AF331522; CAC88370.1; -; mRNA.
DR EMBL; BC094730; AAH94730.1; -; mRNA.
DR EMBL; ENSG00000147606; Homo sapiens.
DR Ensembl; ENSG00000147606; Homo sapiens.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR011547; Sulph_transp.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR PROSITE; PS0801; STAS; 1.
SQ SEQUENCE 656 AA; 72213 MW; 9F8D2E140B5C00CB CRC64;

Query Match 100.0%; Score 2232; DB 2; Length 656;
Best Local Similarity 100.0%; Pred. No. 4.5e-156;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTGAKKKKMLSKMHTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAVQ 60
Db 1 MTGAKKKKMLSKMHTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAVQ 60

Qy 61 QVTQGLAPAVLSSVHPVFGLYGSLFPALIIYAFGMGHVATGTFTALTSLISANAVERIVP 120
Db 61 QVTQGLAPAVLSSVHPVFGLYGSLFPALIIYAFGMGHVATGTFTALTSLISANAVERIVP 120

Qy 121 QNMQLTTQSTSVLGLSDFEMQRIHVAAAVSFLGGVIQVAMFVLQGSATFVVTPEVIS 180
Db 121 QNMQLTTQSTSVLGLSDFEMQRIHVAAAVSFLGGVIQVAMFVLQGSATFVVTPEVIS 180

Qy 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYAVFENIKSVRLAALLLSLISIV 240
Db 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYAVFENIKSVRLAALLLSLISIV 240

Qy 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVGVGHIPQGI PSRAP 300
Db 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVGVGHIPQGI PSRAP 300

Qy 301 PMNLSAVITEAFGVALVGYVASLALAQSAKPKYISIDNQEFLAHGLSNIVSFPFCI 360
Db 301 PMNLSAVITEAFGVALVGYVASLALAQSAKPKYISIDNQEFLAHGLSNIVSFPFCI 360

Qy 361 PSAAAMGRTAGLYSTGAKTQVACLISCFVLIIYVIAIGPLLYWLPVCVLASIIIVVGLKGM 420
Db 361 PSAAAMGRTAGLYSTGAKTQVACLISCFVLIIYVIAIGPLLYWLPVCVLASIIIVVGLKGM 420

Qy 421 LIQFRDLKKYNNVDKIDWG 439
Db 421 LIQFRDLKKYNNVDKIDWG 439

RESULT 2

Q8TE53 HUMAN
ID Q8TE53 HUMAN PRELIMINARY; PRT; 663 AA.
AC Q8TE53;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Anion transporter.
GN Name=SLC26A7; Synonyms=SLC26A7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;

RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Kidney;
RC TISSUE=Kidney;
RX MEBLINE=21818883; PubMed=11829495; DOI=10.1006/geno.2002.6689;
RA Vincourt J.B., Jullien D., Kossida S., Amalric F., Girard J.P.;
RT "Molecular cloning of SLC26A7, a novel member of the SLC26
RT sulfate/anion transporter family, from high endothelial venules and
RT kidney.",
RL Genomics 79:249-256 (2002).

RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Kidney;
RC TISSUE=Kidney;
RA Vincourt J., Jullien D.L., Amalric F., Girard J.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ413230; CAC88372.1; -; mRNA.
DR Ensembl; ENSG00000147606; Homo sapiens.
DR HGNC; HGNC:14467; SLC26A7.
DR GO; GO:0016021; C:integral
DR InterPro; IPR002645; STAS.
DR InterPro; IPR011547; Sulph_transp.
DR Pfam; PF00916; Sulfate_transp; 1.
DR PROSITE; PS0801; STAS; 1.
SQ SEQUENCE 663 AA; 73032 MW; 1F8967404F54DC9F CRC64;

Query Match 100.0%; Score 2232; DB 2; Length 663;
Best Local Similarity 100.0%; Pred. No. 4.5e-156;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTGAKKKKMLSKMHTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAVQ 60
Db 1 MTGAKKKKMLSKMHTPOCEDIIQWCRRLPILDWAPHYLNKENLLPDTVSGIMLAVQ 60

Qy 61 QVTQGLAPAVLSSVHPVFGLYGSLFPALIIYAFGMGHVATGTFTALTSLISANAVERIVP 120
Db 61 QVTQGLAPAVLSSVHPVFGLYGSLFPALIIYAFGMGHVATGTFTALTSLISANAVERIVP 120

Qy 121 QNMQLTTQSTSVLGLSDFEMQRIHVAAAVSFLGGVIQVAMFVLQGSATFVVTPEVIS 180
Db 121 QNMQLTTQSTSVLGLSDFEMQRIHVAAAVSFLGGVIQVAMFVLQGSATFVVTPEVIS 180

Qy 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYAVFENIKSVRLAALLLSLISIV 240
Db 181 AMTTGAATHVVTQVKYLLGKMPYISGPLGFFIYAVFENIKSVRLAALLLSLISIV 240

Qy 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVGVGHIPQGI PSRAP 300
Db 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNMNTYGLVGVGHIPQGI PSRAP 300

Qy 301 PMNLSAVITEAFGVALVGYVASLALAQSAKPKYISIDNQEFLAHGLSNIVSFPFCI 360
Db 301 PMNLSAVITEAFGVALVGYVASLALAQSAKPKYISIDNQEFLAHGLSNIVSFPFCI 360

Qy 361 PSAAAMGRTAGLYSTGAKTQVACLISCFVLIIYVIAIGPLLYWLPVCVLASIIIVVGLKGM 420
Db 361 PSAAAMGRTAGLYSTGAKTQVACLISCFVLIIYVIAIGPLLYWLPVCVLASIIIVVGLKGM 420

Qy 421 LIQFRDLKKYNNVDKIDWG 439
Db 421 LIQFRDLKKYNNVDKIDWG 439

RESULT 3

Q8TE54 HUMAN
ID Q8TE54 HUMAN PRELIMINARY; PRT; 656 AA.
AC Q8TE54;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Anion transporter.
GN Name=SLC26A7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2006, 18:40:15 ; Search time 188 Seconds
(without alignments)
1025.996 Million cell updates/sec

Title: US-10-761-033-2_COPY_1_439

Perfect score: 2232

Sequence: 1 MTGAKRKXKMLNSKMTTPQ.....MLIQFDLKKYNNVDKIDWG 439

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:.*
1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*
9: geneseqp2005s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|----------|--------------------|
| 1 | 2232 | 100.0 | 598 | AAU74917 | Aau74917 Novel hum |
| 2 | 2232 | 100.0 | 605 | AAU74915 | Aau74915 Novel hum |
| 3 | 2232 | 100.0 | 621 | AAU74913 | Aau74913 Novel hum |
| 4 | 2232 | 100.0 | 656 | AAU74916 | Aau74916 Novel hum |
| 5 | 2232 | 100.0 | 663 | AAU74914 | Aau74914 Novel hum |
| 6 | 2232 | 100.0 | 663 | ADC35057 | Adc35057 Human SUT |
| 7 | 2232 | 100.0 | 679 | AAU74912 | Aau74912 Novel hum |
| 8 | 2227 | 99.8 | 656 | AAE22123 | Aae22123 Human ani |
| 9 | 2227 | 99.8 | 656 | ADC35056 | Adc35056 Human SUT |
| 10 | 2077 | 93.1 | 656 | ABR84485 | Abr84485 Human ani |
| 11 | 2077 | 93.1 | 656 | ABR84486 | Abr84486 Mouse ani |
| 12 | 2065.5 | 92.5 | 671 | ABM82599 | Abm82599 Human dia |
| 13 | 2064.5 | 92.5 | 671 | ABG61534 | Abg61534 Human tra |
| 14 | 938.5 | 42.0 | 431 | ABG06299 | Abg06299 Novel hum |
| 15 | 825.5 | 37.0 | 555 | ABB99660 | Abb99660 Amino aci |
| 16 | 805 | 36.1 | 735 | ADM56357 | Adm56357 Frog SLC2 |
| 17 | 774 | 34.7 | 739 | ADM56380 | Adm56380 Mouse SLC |
| 18 | 759 | 34.0 | 739 | AAV08929 | Aav08929 Rat DTBST |
| 19 | 759 | 34.0 | 739 | ADD46847 | Add46847 Rat Prote |
| 20 | 753.5 | 33.8 | 739 | AAV08928 | Aav08928 Human DTD |
| 21 | 753.5 | 33.8 | 739 | AAU74765 | Aau74765 Human SLC |
| 22 | 753.5 | 33.8 | 739 | AAU74763 | Aau74763 Human SLC |
| 23 | 753.5 | 33.8 | 739 | AAU74762 | Aau74762 Human SLC |
| 24 | 753.5 | 33.8 | 739 | AAU74764 | Aau74764 Human SLC |

ALIGNMENTS

RESULT 1

AAU74917
ID AAU74917 standard; protein; 598 AA.
XX
AC AAU74917;
XX
DT 23-APR-2002 (first entry)
XX
DE Novel human protein (NHP) sequence #6.
XX
KW Novel human protein; NHP; transporter protein; polymorphism;
KW mental disorder; biological disease; medical disorder.
XX
OS Homo sapiens.
XX
FN WO200194583-A2.
XX
PD 13-DEC-2001.
XX
PF 06-JUN-2001; 2001WO-US018393.
XX
PR 07-JUN-2000; 2000US-0210045P.
XX

(LEXI-) LEXICON GENETICS INC.

Walke DW, Scoville J;
WPI; 2002-147673/19.

N-PSDB; ABK12981.
XX

New human polynucleotides encoding proteins that share sequence similarity with mammalian transporter proteins, useful for determining genomic structures, identifying polymorphisms, or as reagents for diagnosis or drug screening.

Claim 1; Page 46-47; 48pp; English.

The present invention relates to new isolated nucleic acid molecules comprising a nucleotide sequence encoding a protein that shares sequence similarity with mammalian transporter proteins. The invention also relates to a nucleotide sequence that hybridises under stringent conditions to the nucleotide sequence comprising 1971 bp (ABK12980) fully defined in the specification or its complement. The isolated nucleic acid and the protein it encodes are useful for identifying a coding sequence and mapping a unique gene to a particular chromosome, identifying and characterising the temporal and tissue specific expression of a gene, screening a human genomic library, determining the genomic structure of a

CC given locus/allele or designing diagnostic tests. The nucleic acids and
CC proteins are particularly useful for identifying polymorphisms and in
CC amplification assays to detect mutations within the exons, introns and
CC splice sites that can be used in diagnostics and pharmacogenomics. These
CC are also useful for generating antibodies, as reagents in diagnostic
CC assays or for identifying other cellular gene products related to novel
CC human proteins. The nucleotides can be used as reagents in assays for
CC screening for compounds that can be employed as pharmaceutical reagents
CC useful in the therapeutic treatment of mental, biological or medical
CC disorders and diseases. The present amino acid sequence represents novel
CC human protein #6 that is one of several (AAU74912- AAU74917) novel human
CC proteins (NHP) of the invention
XX
SQ Sequence 598 AA;

Query Match 100.0%; Score 2232; DB 5; Length 598;
Best Local Similarity 100.0%; Pred. No. 5.5e-226;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTGAKRKKKSLWSKMHPTQCEDIIQWCRRLPILDWAPHYNLKENLLPDTVSGIMLAVQ 60
Db 1 MTGAKRKKKSLWSKMHPTQCEDIIQWCRRLPILDWAPHYNLKENLLPDTVSGIMLAVQ 60
Qy 61 QVTQGLAFVAVLSSVHPVGLYGSLLFPALIIYAFMGHHVATGTFTALTSLSANAVERTVP 120
Db 61 QVTQGLAFVAVLSSVHPVGLYGSLLFPALIIYAFMGHHVATGTFTALTSLSANAVERTVP 120
Qy 121 QNMQLTTQSNVTSVLGLSDFEMQRIHVAAAVSFLGGVIVQAMFVLQGSATFVVTPEPVS 180
Db 121 QNMQLTTQSNVTSVLGLSDFEMQRIHVAAAVSFLGGVIVQAMFVLQGSATFVVTPEPVS 180
Qy 181 AMTTGAATHVVTQVKYLLGKMPYISGLPGFFIYIYAVFENIKSVRLLEALLLSISIV 240
Db 181 AMTTGAATHVVTQVKYLLGKMPYISGLPGFFIYIYAVFENIKSVRLLEALLLSISIV 240
Qy 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNNNTYGLVGVHIGPQIPSPRAP 300
Db 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNNNTYGLVGVHIGPQIPSPRAP 300
Qy 421 LIQFRDLKKYWNVDKIDWG 439
Db 421 LIQFRDLKKYWNVDKIDWG 439

RESULT 2
AAU74915
ID AAU74915 standard; protein; 605 AA.

XX AC AAU74915;
XX
DT 23-APR-2002 (first entry)
DE
XX Novel human protein (NHP) sequence #4.
XX
KW Novel human protein; NHP; transporter protein; polymorphism;
XX mental disorder; biological disease; medical disorder.

OS Homo sapiens.
XX
FN WO200194583-A2.
XX
PD 13-DEC-2001.
XX
PF 06-JUN-2001; 2001WO-US018393.

PR 07-JUN-2000; 2000US-0210045P.
XX (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Scoville J;
XX WPI; 2002-147673/19.
DR N-PSDB; ABK12979.
XX
PT New human polynucleotides encoding proteins that share sequence
PT similarity with mammalian transporter proteins, useful for determining
PT genomic structures, identifying polymorphisms, or as reagents for
PT diagnosis or drug screening.
XX
PS Claim 1; Page 42-43; 48pp; English.
XX
CC The present invention relates to new isolated nucleic acid molecules
CC comprising a nucleotide sequence encoding a protein that shares sequence
CC similarity with mammalian transporter proteins. The invention also
CC relates to a nucleotide sequence that hybridizes under stringent
CC conditions to the nucleotide sequence comprising 1971 bp (ABK12980) fully
CC defined in the specification or its complement. The isolated nucleic acid
CC and the protein it encodes are useful for identifying a coding sequence
CC and mapping a unique gene to a particular chromosome, identifying and
CC characterising the temporal and tissue specific expression of a gene,
CC screening a human genomic library, determining the genomic structure of a
CC given locus/allele or designing diagnostic tests. The nucleic acids and
CC proteins are particularly useful for identifying polymorphisms and in
CC amplification assays to detect mutations within the exons, introns and
CC splice sites that can be used in diagnostics and pharmacogenomics. These
CC are also useful for generating antibodies, as reagents in diagnostic
CC assays or for identifying other cellular gene products related to novel
CC human proteins. The nucleotides can be used as reagents in assays for
CC screening for compounds that can be employed as pharmaceutical reagents
CC useful in the therapeutic treatment of mental, biological or medical
CC disorders and diseases. The present amino acid sequence represents novel
CC human protein #4 that is one of several (AAU74912- AAU74917) novel human
CC proteins (NHP) of the invention
XX
SQ Sequence 605 AA;

Query Match 100.0%; Score 2232; DB 5; Length 605;
Best Local Similarity 100.0%; Pred. No. 5.6e-226;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTGAKRKKKSLWSKMHPTQCEDIIQWCRRLPILDWAPHYNLKENLLPDTVSGIMLAVQ 60
Db 1 MTGAKRKKKSLWSKMHPTQCEDIIQWCRRLPILDWAPHYNLKENLLPDTVSGIMLAVQ 60
Qy 61 QVTQGLAFVAVLSSVHPVGLYGSLLFPALIIYAFMGHHVATGTFTALTSLSANAVERTVP 120
Db 61 QVTQGLAFVAVLSSVHPVGLYGSLLFPALIIYAFMGHHVATGTFTALTSLSANAVERTVP 120
Qy 121 QNMQLTTQSNVTSVLGLSDFEMQRIHVAAAVSFLGGVIVQAMFVLQGSATFVVTPEPVS 180
Db 121 QNMQLTTQSNVTSVLGLSDFEMQRIHVAAAVSFLGGVIVQAMFVLQGSATFVVTPEPVS 180
Qy 181 AMTTGAATHVVTQVKYLLGKMPYISGLPGFFIYIYAVFENIKSVRLLEALLLSISIV 240
Db 181 AMTTGAATHVVTQVKYLLGKMPYISGLPGFFIYIYAVFENIKSVRLLEALLLSISIV 240
Qy 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNNNTYGLVGVHIGPQIPSPRAP 300
Db 241 LVLVKELNEQPKRKIKVLPVDLVLIIAASFCYCTNNNTYGLVGVHIGPQIPSPRAP 300
Qy 301 PMNLSAVITEAFGVAVLGVVAVSLAQGSAAKPKYISIDNQEFLAHLGSLNIVSFPFCI 360
Db 301 PMNLSAVITEAFGVAVLGVVAVSLAQGSAAKPKYISIDNQEFLAHLGSLNIVSFPFCI 360
Qy 361 PSAAAMGRTAGLYSTGAKTQVACILISCFVLIVYIYAGPLLYWLPVCVLASIIIVVGLKGM 420
Db 361 PSAAAMGRTAGLYSTGAKTQVACILISCFVLIVYIYAGPLLYWLPVCVLASIIIVVGLKGM 420